UNIVERSIDADE DE LISBOA
FACULDADE DE CIÊNCIAS
DEPARTAMENTO DE FÍSICA

Enhancing reprogramming and transdifferentiation through long non-coding RNAs

Miguel Torres Santana

Mestrado Integrado em Engenharia Biomédica e Biofísica

Perfil de Engenharia Clínica e Instrumentação Médica

Dissertação orientada por:
Doutor Bruno Miguel Bernardes de Jesus
Professor Doutor Pedro Miguel Dinis de Almeida

[2017]

UNIVERSIDADE DE LISBOA
FACULDADE DE CIÊNCIAS
DEPARTAMENTO DE FÍSICA

Enhancing reprogramming and transdifferentiation through long non-coding RNAs

Miguel Torres Santana

Mestrado Integrado em Engenharia Biomédica e Biofísica

Perfil de Engenharia Clínica e Instrumentação Médica

Dissertação orientada por:
Doutor Bruno Miguel Bernardes de Jesus
Professor Doutor Pedro Miguel Dinis de Almeida

[2017]

This page intentionally left blank

Agradecimentos

Quero expressar em primeiro lugar o meu agradecimento ao Doutor Bruno Jesus, pela orientação, ajuda e apoio constante ao longo deste último ano letivo e pela oportunidade que me proporcionou de participar neste projeto.

Agradeço ainda ao António Franco por me ter acompanhado, ajudado e ensinado no período de adaptação e ao Sérgio Marinho pela sua constante boa disposição e disponibilidade e boa vontade em me ajudar.

Agradeço ainda a todos os restantes colegas e amigos do laboratório e do IMM, pela companhia, apoio constante, boa disposição e amizade.

Agradeço também à minha família por todo o apoio, aconselhamento e por estarem sempre ao meu lado, mesmo estando distantes.

Por fim, não posso deixar de agradecer a todos os meus amigos que me acompanharam e ajudaram ao longo destes últimos anos.

Resumo

Foi recentemente desenvolvido um novo método revolucionário capaz de reprogramar fibroblastos em células pluripotentes induzidas através da expressão de 4 fatores de transcrição (*Oct4*, *Sox2*, *Klf4* e *c-Myc*). A reprogramação de fibroblastos em células pluripotentes induzidas (*iPSC*) foi um grande avanço científico com possíveis aplicações clínicas e fins terapêuticos, no entanto, à medida que as células diferenciadas (células somáticas) vão envelhecendo devido à acumulação de marcas genéticas e epigenéticas, estas tornam-se mais resistentes à sua conversão para um estado pluripotente.

Posto isto, um dos principais objetivos deste projeto passava por perceber o impacto que o envelhecimento tem na reprogramação de células humanas em células estaminais pluripotentes induzidas humanas (*hiPSCs*). De facto, ao reprogramar fibroblastos adultos de ratinho em células pluripotentes induzidas (*miPSCs – mouse induced pluripotent stem cells*), observou-se que o envelhecimento celular estava a atuar como uma barreira, reduzindo a eficiência da reprogramação celular. Curiosamente, nenhuma correlação entre a eficiência da reprogramação celular e o envelhecimento foi observada na reprogramação de células humanas. Ao realizar a reprogramação celular de fibroblastos embrionários humanos (WI38) e fibroblastos humanos com 3 anos (3yr) com baixa passagem (passage 4), observou-se o mesmo número de células *hiPSCs* geradas. Sugerindo assim, e ao contrário do esperado e observado em células de ratinho, que a idade das células humanas utilizadas para formar *hiPSCs* não dificulta de forma significativa a reprogramação celular.

Contudo, foi observado que o número de passagens das células em cultura tinha uma contribuição importante na eficiência da reprogramação celular de fibroblastos humanos em *hiPSCs*. Ao tentar reprogramar fibroblastos humanos com uma baixa passagem (passagem 4) e uma alta passagem (passagem 7), e apesar de as células com passagem 7 expressarem níveis mais elevados de hOCt4, apenas as células humanas com uma baixa passagem reprogramaram. Estes resultados sugerem que o número de passagens celulares tem uma contribuição importante na eficiência da reprogramação celular.

De facto, estes resultados podem ser explicados pela simples razão de que cada passagem celular realizada em cultura, aumenta o risco de ocorrer dano no ADN, mutações e ainda alterações nas características celulares originando assim, alterações na morfologia, na resposta a estímulos, na taxa de crescimento, na expressão de proteínas e na eficiência da transfecção celular. De acordo com Leonard Hayflick e Paul Moorhead, as células humanas têm um número limitado de divisões celulares em cultura que poderá variar entre tipos de células. Cada divisão celular pode induzir um encurtamento dos telómeros que poderá resultar em senescência. Podendo esta ser ainda induzida através de dano molecular que ocorre de forma aleatória, pelo stress oxidativo e pela danificação do ADN. A senescência pode ainda ser acumulada em alguns tecidos contribuindo para a disfunção orgânica. Sugere-se portanto, que devido ao número elevado de passagens celulares, os fibroblastos embrionários humanos e os fibroblastos humanos com 3 anos sofreram o encurtamento dos seus telomeros, levando assim, a um fenótipo senescente, reduzindo a eficiência do processo de reprogramação celular em células humanas.

Um outro objetivo deste projeto passava por encontrar outras estratégias celulares que ajudem a ultrapassar a limitação do envelhecimento na reprogramação celular de células humanas, através da modulação de RNAs longos não codificantes (*lncRNA*). No entanto, como os resultados referentes ao envelhecimento demonstraram uma não influência na eficiência da reprogramação de células humanas em *hiPSCs*, decidiu-se desvendar e entender, qual a função do *lncRNA* Zeb2NAT na reprogramação celular e qualidade de células humanas. Embora, já reportado anteriormente pelo nosso laboratório,

que a supressão do *lncRNA* Zeb2NAT em células de ratinho, utilizando oligonucleótidos contra-senso (*anti sense*), denominados por *LNAs* aumentam significativamente a reprogramação celular de fibroblastos de ratinho envelhecidos ajudando assim contornar as barreiras mesenquimais, ainda nada se sabia sobre o impacto que a diminuição do Zeb2 e do Zeb2NAT poderia ter na reprogramação celular de células humanas.

Contudo, utilizando a mesma abordagem, acima descrita, em fibroblastos embrionários humanos e em fibroblastos humanos com 3 anos, verificou-se um atraso na formação de *hiPSCs* das células humanas que sofreram uma desregulação tanto do Zeb2 como do Zeb2NAT, comparativamente às células do controlo e do *LNA-controlo* (um LNA não específico a nenhuma sequência genómica humana, utilizado como um controlo negativo da transfeção de *LNAs*). De facto, 14 dias após a primeira transfeção com *LNAs* em ambas as duas linhas celulares, apenas começaram a formar-se *hiPSCs* no controlo e no *LNA-controlo*, tendo o controlo um número maior de hiPSCs formadas em ambas as células, comparativamente à condição com *LNA-controlo*. No entanto, 19 dias e 35 dias depois da primeira transfeção, as WI38 com a desregulação do Zeb2 e do Zeb2NAT, respetivamente, começaram a reprogramar. Nenhuma formação de hiPSCs nas células humanas de 3 anos com a desregulação do Zeb2 e do Zeb2NAT foi observada.

Tendo em consideração que as células de ratinho após diminuição dos níveis do Zeb2 e do Zeb2-NAT começaram a reprogramar de forma mais eficiente que o controlo e o *LNA-controlo*, estes resultados em células humanas sugerem que a desregulação do Zeb2 e do Zeb2NAT pode estar, de certa maneira, a atrasar e até mesmo a atuar como um bloqueador da reprogramação celular humana.

Este bloqueio/atraso que se observou na reprogramação de fibroblastos humanos em *hiPSCs* poderia ter como principal responsável a transfeção de *LNAs*, necessária para diminuir a expressão do Zeb2 e do Zeb2NAT. De facto, já foi demonstrado que o uso do reagente de transfeção *Lipofectamine* ativa algum stress nos genes afetando o ciclo da regulação e/ou a sinalização metabólica nas células. No entanto, as células com a condição do *LNA-controlo* reprogramaram com a mesma rapidez que o controlo (sem LNAs). E a diferença do número de *hiPSCs* geradas pelo *LNA-controlo*, comparativamente ao controlo, acaba por não ser significativamente diferente. Isto sugere, que embora a transfeção de *LNAs* tenha um impacto na reprogramação celular, diminuindo a eficiência desta, existe um outro fator que está a contribuir para que as células com a diminuição da expressão do Zeb2 e do Zeb2-NAT tenham uma redução na eficiência da reprogramação celular em células humanas.

Todavia, ainda é incerto a razão de a desregulação do Zeb2 e do Zeb2NAT terem atrasado e até mesmo bloqueado a reprogramação celular em células humanas. É possível que possa ser devido ao mecanismo que rege o Zeb2 que, de alguma forma, é diferente comparativamente ao mecanismo observado em ratinhos. Outra possibilidade a ter em conta é a necessidade de realizar uma otimização ao protocolo da transfeção de *LNAs*, de forma a adaptar esta ao protocolo da reprogramação celular.

Apesar disto, acreditamos que esta abordagem constitui uma nova estratégia para estudar o impacto dos *lncRNAs* na reprogramação celular e antecipamos ainda que os resultados produzidos irão gerar contribuições importantes na área de investigação do envelhecimento e da reprogramação celular.

Palavras-chave: Reprogramação celular, fibroblastos humanos, células estaminais pluripotentes induzidas humanas, envelhecimento, Zeb2NAT

Abstract

Revolutionary progress has been achieved recently following the discovery of cellular reprogramming by the expression of a combination of 4 transcription factors (Oct4, Sox2, Klf4 and c-Myc). The reprogramming of fibroblasts to induced pluripotent stem cells (iPSC) was a major scientific advance however, as differentiated cells grow old and due to the accumulation of genetic and epigenetic marks, they become more resistant to be converted back to a pluripotent state.

In fact, when tried to reprogram mice cells into miPSCs, aging was acting as a barrier, reducing the efficiency of reprogramming aged cells. However, through our recent findings no correlation between cellular reprogramming efficiency and aging was found in human cells. However, we observed that the number of passages had an important contribution in the efficiency of reprogramming human fibroblasts into hiPSCs. Reprogramming experiments with human fibroblasts with lower passage (passage 4) and high passage (passage 7) show that only the human cells with a lower passage, reprogrammed. The higher the number of passages in vitro, the lower the efficiency of cellular reprogramming of human cells.

The other main interest of this project converged on cellular strategies to overcome this aging limitation by modulating long noncoding RNAs (lncRNAs). However, as the results showed a non-influence of aging on reprogramming human cells into hiPSCs it was decided to focus on understanding exactly the role of the lncRNA Zeb2NAT on cellular reprogramming of human cells. As previously reported by our lab, the suppression of Zeb2NAT in mice cells significantly increased the reprogramming of aged fibroblasts. However, when tested the same approach in human cells, a delay in generating hiPSCs was observed, suggesting that the knockdown of Zeb2 and Zeb2NAT can be acting as a blocker of cellular reprogramming in human cells, going against what was observed in mice cells.

Still, it's uncertain why downregulation of Zeb2 and Zeb2NAT were delaying and blocking the cellular reprogramming in human cells. It is possible that the mechanisms of Zeb2 in humans are, somehow, different from those observed in mice. Another possibility to take in account can be due the fact the protocol isn´t fully optimized.

Despite this, we believe this approach constitutes a novel strategy to study the impact of lncRNAs in cellular reprogramming and we anticipate that the output of this proposal will generate important contributions to the aging and cellular reprogramming research field.

Keywords: Cellular reprogramming, human fibroblasts, human iPSCs, Aging, lncRNA Zeb2NAT

Table of contents

List of tables

List of figures

List of abbreviations

3yr – 3yr human fibroblasts

293T – Human embryonic kidney 293T cells

AP – Alkaline phosphatase

cDNA – Complementary DNA

c-Myc – MYC

CIN – Chromosomal instability

DMEM – Dulbecco's Modified Eagle Medium

E14 – Mouse stem cells E14

ECATs – ES cell-associated transcripts

ESCs – Embryonic stem cells

EMT – Epithelial-mesenchymal transition

FBS – Fetal Bovine Serum

HGPS – Hutchinson-Gilford Progeria syndrome

hiPSCs – Human induced pluripotent stem cells

GATA1 – Gata-binding protein 1

GFP – Green fluorescence protein

hOct4 – Human Oct4

ICM – Inner cell mass

iPSCs – Induced Pluripotent Stem Cells

IRES – Internal ribosome entry site

KLF4 - Krüppel-like factor 4

KSR – KnockOut™ Serum

LIF – Leukemia inhibitory factor

LNA – Locked nucleic acid

lncRNA-RoR – Long non-coding RNAs regulator of reprogramming

lncRNAs – Long non-coding RNAs

MEFs – Mouse embryonic fibroblasts

mESCs – Mouse embryonic stem cells

MET – Mesenchymal-epithelial transition

miRNAs – Micro RNAs

miPSCs – mouse induced pluripotent stem cells

mRNAs – Messenger RNAs

MYOD – Myoblast determination protein

ncRNAs – Non-coding RNAs

NEAA – Non-essential amino acids solution

SCNT – Somatic cell nuclear transfer

snoRNPs – Small nucleolar ribonucleoproteins

snRNPs – Small nuclear ribonucleoproteins

Oct4 - Octamer-binding transcription factor 4

PBS – Phosphate Buffered Saline

PCR – Polymerase chain reaction

Pen/Strep – Penicillin Streptomycin

PSCs – Pluripotent stem cells

PUF – Pumilio-Fem3-binding factor family

PUM1 – PUMILIO 1

PUM2 – PUMILIO 2

RT – Room temperature

RT-qPCR – Real-time quantitative PCR

Sox2 - SRY-box 2

WI38 – human embryonic fibroblasts

This page intentionally left blank

1. Introduction and literature review

1.1. Stem cells

The increase in life expectancy lead to an increase in the number of people affected by some kind of condition in particular associated to age-related diseases such as cancer or heart conditions. Due to this, different areas of science are being developed to find a faster and better response to age-related problems. One example is the stem cells field. Stem cells are one of the most promising fields for regenerative medicine and clinical application having an enormous therapeutic potential to replace damaged tissues and cells[1].

One of the first studies identifying stem cells was performed in 1981 by Evans and Kaufman and Gail R. Martin describing independently how to successfully derive mouse embryonic stem cells (mESCs) from mouse blastocysts. These were seminal works on the stem cell field, demonstrating the capability to derive and expand stem cells *in vitro*[2].

A stem cell has to satisfy three essential criteria. First, a stem cell is a cell not yet specialized for any particular function. The second, it must has the ability to differentiate into a specialized cell type under the right conditions, known as cell plasticity, and the third one argues that a stem cell must divide (self-renew) indefinitely. The term self-renewing refers to the ability to undergo multiple divisions while maintaining an undifferentiated state[3].

Stem cells can be classified into two main categories based on their self-renewing capacity and plasticity, namely "embryonic stem cells" and adult or somatic stem cells (i.e. "non-embryonic")[2]. Embryonic stem cells (ESCs) are a pluripotent cell type, meaning that they have the capacity to differentiate into the three germ layers (mesoderm, endoderm and ectoderm). Due to this capacity of pluripotency, ESCs are considered to have the greatest potential for regenerating new tissues for patient treatment[2].

An adult stem cell is a type of cell that meets all the three of the above criteria but is found within already specialized human tissue. It's also typically multipotent presenting the capacity to differentiate only into their corresponding tissues or organs to replace the injured cells. For instance, a hematopoietic stem cell can differentiate only into blood cells due to its lineage restriction[4].

The presence of stem cells in adult tissues provides a chance of self-renewing after some trauma or natural cell death. This can be pertinent for many tissues such as the hair, skin, bone marrow, central nervous system, blood, liver, kidney or male germ cells[3].

The three different primary germ layers that ESCs can differentiate into are: The mesoderm (middle layer), endoderm (internal layer) and the ectoderm (external layer) (Figure 1.1). Each layer has different types of cells corresponding to its specificity. For instance, alveolar cells, thyroid cells and pancreatic cells are from the endoderm layer while red blood cells and skin cells are from the mesoderm and ectoderm, correspondingly. The adult tissue-specific multipotent cells can only differentiate into one of this three embryonic germ layers[5].

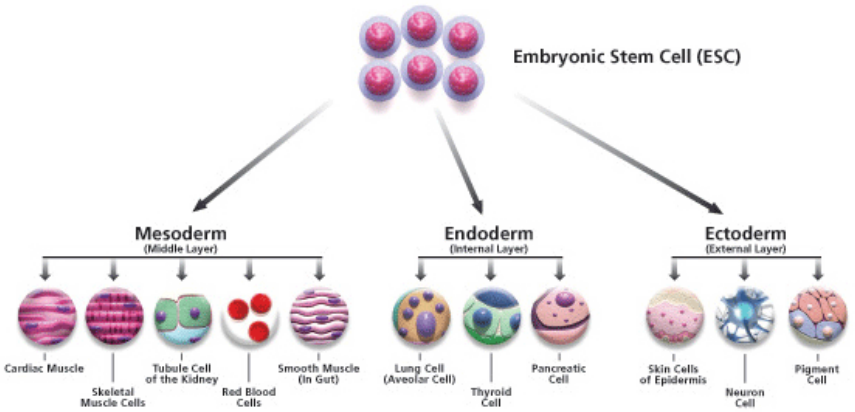


Figure 1.1 - Pluripotent stem cell differentiation into all three germ layers (Mesoderm, Endoderm and Ectoderm). Adapted from Juty N. et all, 2008.

1.2. Cellular reprogramming

Due to ethical limitations to acquire human embryonic stem cells for medical treatments (since it involves the manipulation of human blastocysts) new methodologies to obtain embryonic stem cells (ESC) were a request in the stem cell field[2].

One of the first experiments to obtain in vitro ESC was reported in 1962 by Sir John Gurdon, who reported a method called somatic cell nuclear transfer (SCNT). During SCNT, the nucleus of a somatic cell is transferred to an enucleated and unfertilized egg of the same species. After some divisions of the egg a blastocyst is formed, an early stage embryo that is genetically identical to the donor of the somatic cell is generated. Through the isolation of inner cell mass (ICM) of the blastocyst by immunosurgery was possible to create a culture of embryonic stem cells[6,7]. This experiment demonstrated that the nuclei of a somatic cell maintain all genetic being possible to reprogram a somatic cell to an embryonic, pluripotent state by experimental manipulation[8]. Using this method was born the famous Dolly the sheep, the first mammal cloned from an adult somatic cell[9].

Later, several studies revealed that the profile of gene expression in somatic cells can be changed through fusion with other cell types, thus causing reprogramming of these cells. In 1983, Helen M. Blau showed that silence of muscle-specific genes in human amniocytes is activated after cell fusion with mouse muscle cells (generating what is called of heterokaryons). It was also demonstrated, in the same year, that inactivated X chromosomes in female somatic cells, such as thymocytes or bone marrow cells, could be reactivated by fusion with teratocarcinoma-derived cells in which both X chromosomes were activated. Others reported that somatic cells could be reprogrammed to express genes that are predominantly expressed in pluripotent cells in vitro and/or in vivo, such as Oct4, by fusing them with pluripotent cells (for example, ES cells)[10,11]. This suggest that pluripotent stem cells (PSCs) have the potential to reprogram somatic cells toward pluripotency, suggesting the existence of one or more reprogramming factors that can "erase" the "memories" of somatic cells[8,12].

In 2006, Shinya Yamanaka lab, described a protocol to induce pluripotency in already differentiated cells by "reprogramming" them through the expression of four transcription factors, Octamer-binding transcription factor 4 (Oct4), SRY-box 2 (Sox2), Krüppel-like factor 4 (KLF4) and MYC (c-MYC), known has OSKM (Figure 1.2) [13,14].

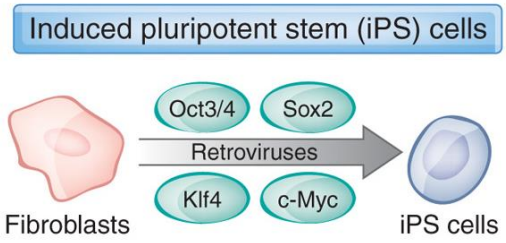


Figure 1.2 - iPSC reprogramming protocol by introducing genes encoding four transcription factors[15].

These cells were called induced pluripotent stem cells (iPSC) and have the same morphology, growth properties and markers of ESCs. The process to induce pluripotent stem cells from already differentiated cells, using transcription factors, is called cellular reprograming[13,14].

The existence of reprogramming factors was also demonstrated for direct fate conversion of mammalian cells. A method that allow to convert directly somatic cells into different specialized cells without a pluripotent state transition[16]. For instance, myoblast determination protein (MYOD), alone was sufficient to transform mouse fibroblasts to myoblast. It was also demonstrated that the ectopic expression of erythroid transcription factor GATA-binding protein 1 (GATA1) could convert myeloblasts to megakaryocyte and erythrocyte precursors. The scientific term given to this process of converting somatic cells into a different somatic lineage is known as transdifferentiation[8,17].

Yamanaka S. and Takahashi K. started to analyze mouse ES cells to identify genes underlying characteristics, such as pluripotency and proliferation. This investigation led to the identification of ES cell-specific genes, referred to as ES cell-associated transcripts (ECATs). After various experiments, including the generation of knockout ES cells and knockout mice, they discovered that NANOG is an ECAT being crucial for the maintenance of pluripotency in both ES cells and early embryos. It was also uncovered that the overexpression of NANOG allowed mouse ES cells to self-renew, even without the presence of leukemia inhibitory factor (LIF), which is an essential cytokine for the maintenance of mouse cell pluripotency in serum containing medium. Some others transcription factors of pluripotency were also discovered, such as KLF4, c-MYC, Sox2 and Oct4, as mentioned before[8].

The transcription factors were tested in several murine and human cell lines. For this, it was initially used a retroviral transduction system, due to his high efficiency in gene delivery, developed by Toshio Kitamura at the University of Tokyo. The cells generated with the use of OSKM factors had ES cell-like properties but displayed somewhat incomplete pluripotency, that is, the characteristics of iPSCs were different to ES cells. This happened because in cellular reprogramming the expression of the four transcription factors (OSKM) should sustain for a sufficient period of time while the cell is being reprogrammed. When the cell is fully reprogrammed, the OSKM expression should be silenced in order to produce iPSCs with the same characteristics of ES cells[8].

Aside of incomplete pluripotency, a crucial issue with iPSCs was the use of retroviral vectors to deliver the reprogramming factors. These vectors integrate into the genome of host cells, potentially causing disruption or aberrant activation of neighboring genes, and pose a risk of reactivation of the reprogramming factors. Indeed, the reactivation of c-MYC induces tumor formation in iPSC-derived chimeric mice. A way to solve this issue was the introduction of efficient, integration-free methods for

cell reprograming, such as adenoviruses, plasmids, transposons, Sendai viruses, synthetic mRNAs and recombinant proteins[8].

Although the pluripotency of fully reprogrammed mouse iPSCs seems to be indistinguishable from that of mouse ES cells, it remains controversial for human iPSCS (hiPSCs), which seems to show differences in gene expression and DNA methylation patterns compared to human ES cells, demonstrating the urge for optimization of human cellular reprogramming[8,18].

1.3. Cellular Aging

Cellular aging is one of most well characterized roadblock for cellular reprogramming. Aging is defined by the functional decline of cells, tissues and organs throughout life. It is also associated with a dramatic increase in a wide range of age-related diseases, including cardiovascular dysfunction, metabolic disorders, neurodegeneration and cancer[19].

Cellular aging is influenced by exposure to extrinsic factors, such as inflammatory cytokines, and intrinsic factors. Old cells accumulate genomic damage, aggregated proteins and display telomere erosion and mitochondrial dysfunction. It's also known that manipulation of some genetic pathways (e.g. insulin-FoxO, Tor, AMPK and Sirtuin) and environmental interventions can delay aging, even if initiated late in life[19].

Through some initial studies that tested how age of the donor of cells affects reprogramming in mice, was suggested that cells from older donors tend to reprogram less efficiently than cells from young mice. For instance, dermal fibroblasts from old mice (>2 years) exhibited a reduction in their ability to generate colonies that stained positive for the stem cell marker alkaline phosphatase (AP) compared to fibroblasts from young adult mice (2 months old) upon expression of the four transcription factors (OSKM) by doxycycline induction. This was also observed in bone marrow cells from old mice (23 months old) that generated 5 times less AP+ colonies than cells from young adults (2 months old)[19]. These studies confirm an age-dependent decline in reprogramming efficiency in mice. Regarding human cells it is still unknown how the aging process affects cellular reprogramming, therefore further tests in human cells are needed to prove if the results observed in mice correlate with the human situation (Table 1.1)[19].

Table 1.1 - Summary of the current literature on the impact of age on reprogramming[19].

Table 1

Summary of the current literature on the impact of age on reprogramming. *Bona fide* iPSCs are defined as iPSCs expressing early and late markers of pluripotency, *in vitro* and/or *in vivo* differentiation potential into the three germ layers

Species	Cell type	Protocol	Age groups	Reprogramming Efficiency	Bona fide iPSCs	Re-differentiation potential	Reference
Mouse	Dermal fibroblasts from ear punches (C57BL/6 mice)	OSKM	2 vs >24 months	2.2-fold higher in younger	Not assessed	Not assessed	Li *et al.*, 2009 [19••]
Mouse	Dermal fibroblasts (B6CBAF1 mice)	OSKM	Juvenile vs 12 months	5-fold higher in younger	Yes	Haematopoietic cells and osteoblasts No age comparison performed	Kim *et al.*, 2010 [21••]
Mouse	Bone marrow cells (C57BL/6 mice)	OSKM	2 vs 23 months	5-fold higher in younger, and twice as fast	Yes	Myeloid cells No age comparison performed	Cheng *et al.*, 2011 [22•]
Mouse	Muscle-derived fibroblasts (C57BL/6-background *mdx* mice)	OSKM	1.5 vs 6 vs 14 months	6-fold higher in 1,5 and 6 months compared to 14 months	Yes	Skeletal muscle lineages No age comparison performed	Wang *et al.*, 2011 [20]
Human	Fibroblasts from various tissues of donors of both sexes and with different disease states	OSKM	8–64 years	No correlation with age	Yes	Definite endoderm. No age comparison performed	Somers *et al.*, 2010 [24]
Human	Dermal fibroblasts	OSKM + NANOG + LIN28	70 years	Not assessed	Yes	Fibroblast No age comparison performed	Suhr *et al.*, 2010 [28]
Human	Dermal fibroblasts of donors of both sexes and with various disease states	OSK and OSKM	29–82 years	Not assessed	Yes	Motor Neurons Efficiency not correlated to age	Boulting *et al.*, 2011 [25]
Human	Dermal fibroblasts	OSKM	84 years	Not assessed	Yes	Neuronal lineages No age comparison performed	Prigione *et al.*, 2011 [30•]
Human	Senescent and proliferative dermal fibroblasts	OSKM + NANOG + LIN28	74–101 years	Not assessed	Yes	Fibroblast No age comparison performed	Lapasset *et al.*, 2011 [26••]
Human	Keratinocytes	OSKM	56–78 years	Not assessed	Yes	Insulin-producing cells	Ohmine *et al.*, 2012 [27]
Human	Dermal fibroblasts	OSKM	106–109 years	Not assessed	Yes	Neural cells No age comparison performed	Yagi *et al.*, 2012 [29•]

One of the mechanisms by which aging could be impairing the reprogramming efficiency in mice is cellular senescence and subsequent loss of division capacity characteristic of aged cells. Senescence is the gradual deterioration of the cellular functions leading to a virtually irreversible cell cycle arrest. Old mice donors contain more populations of cells with senescence or pre-senescence phenotypes thus decreasing reprogramming efficiency. Senescence is also common in human aged samples, although it's not completely known how it affects the reprogramming efficiency of human cells[19].

Reprogramming has the remarkable ability to reverse some cellular and molecular characteristics associated with aging, including cellular senescence, telomere erosion, mitochondrial dysfunction and global changes in gene expression, suggesting that many of the age-associated characteristics that were once thought to be permanent are, in fact, reversible. However the accumulation of nuclear and mitochondrial DNA damage that is associated with aging is unlikely an aspect of aging able to be reversed by reprogramming[19].

1.4. Non-coding RNAs

The vast majority of genomic information is pervasively transcribed into not only a diverse range of protein-coding RNAs but also into long and short non-coding RNAs (ncRNAs). NcRNAs are molecules that are not translated into proteins but have a remarkable variety of biological functions, such as to regulate gene expression at the levels of transcription, RNA processing and translation, or even to protect genomes from foreign nucleic acids, guided DNA synthesis or genome rearrangement[20].

Depending on their function, non-coding RNAs can be classified into the following: microRNAs, piwi-interacting RNAs, small interfering RNAs, long non-coding RNAs, enhancer RNAs and promoter-associated RNAs[20].

Its function could be mediated through the formation of RNA-protein complexes[21]. This RNA-protein complexes occurs when a protein binds to a RNA playing a major role in RNA transcription, RNA processing or protein synthesis[20].

1.4.1. Long non-coding RNAs

Long Non-coding RNAs (lncRNAs) are transcripts with more than 200 nucleotides that don't encode proteins but have an important role in gene expression regulation and protein synthesis, such as, chromatin modification and structure, direct transcriptional regulation, regulation of RNA processing events (splicing, editing, localization, translation and turnover/degradation), facilitation of ribonucleoprotein (RNP) complex formation, gene silencing through production of endogenous small interfering RNAs and regulation of genomic imprinting. lncRNA can also be characterized as one or more of the following models of action: signal, decoy, guide, scaffold and enhancer. They are also usually transcribed by RNA polymerase II and frequently spliced and polyadenylated as other messenger RNAs (mRNAs). Unlike mRNAs, which is abundant and enriched in the cytoplasm, the lncRNAs tends to be expressed at lower levels and are predominantly localized in the nucleus[22].

Several studies of lncRNAs suggested that they can operate through distinct manners, including as signals, scaffolds for protein-protein interactions, molecular decoys, antisense interference and guides to target elements in the genome or transcriptome. They are also involved in phenomena such as imprinting genomic loci, shaping chromosome conformation and allosterically regulating enzymatic activity, but the majority of lncRNAs functions are still unknown. Previous studies also suggested that lncRNAs expression is more cell-type-specific comparing to mRNA expression, meaning that lncRNAs may be involved in cell fate[22].

The ability to regulate gene expression suggests that lncRNAs have influence in the reprogramming of the cells, becoming a viable and useful way to improve cellular reprogramming[23].

An example of an lncRNA candidate with potential role in reprogramming has been described in 2015, by Mao Z. and his team who identified in human fibroblasts 986 down-regulated lncRNAs and 899 up-regulated lncRNAs in senescence cells compared with young cells. Among the lncRNAs, they characterized a senescence-associated lncRNAs, called SALNR that has low expression in senescent cells. When SALNR was overexpressed, cellular senescence was delayed. SALNR interacts with

NF90 (nuclear factor of activated T-cells) which suppress the biogenesis of senescence-associated miRNAs, such as miR-22 and miR181a. The NF90 inhibition results in premature senescence[24].

When human fibroblasts were exposed to Ras-induced (oncogene Ras) stress to activate senescence, NF90 was translocated to nucleolus and couldn't no longer suppress senescence-associated miRNA biogenesis. This was rescued by SALNR overexpression, which antagonized NF90 translocation into nucleolus and rescued its inhibitory activity on senescence-associated miRNA expression. These results suggests that lncRNA SALNR controls cellular senescence through a differential localization of NF90[24].

To understand if lncRNAs could have a role in iPSC or cellular reprogramming, Rinn J. and his colleagues did, in 2010, a comparison between transcriptional profiles of human lncRNAs and protein-coding genes through iPSCs and hESCs derived from human fibroblasts. Gene expression profile analysis revealed that all iPSCs were similar to ESCs[25].

Then, they designed a microarray probing ~900 lncRNAs in the human genome to analyze the expression of lncRNAs in the above cell lines. Through the results they identified 133 lncRNAs that were induced and 104 lncRNAs that were repressed across all iPSCs and ESCs compared with fibroblasts. They identified 28 lncRNAs that had a higher expression in fibroblasts iPSCs relative to ESCs (were referred to as "iPSC-enriched" lncRNAs). With these results they hypothesized that iPSC-enriched lncRNAs have an important role in reprogramming. To test this hypothesis, they profile lncRNA expression in $CD34^+$ hematopoietic stem and progenitor cells, two $CD34^+$ iPSC lines and ESCs. The results correlate with their previous observations, showing that ten of the twenty-eight lncRNAs elevated in fibroblast iPSCs were also elevated in $CD34^+$ iPSCs. Through RT-PCR, eight out of ten iPSC-enriched lncRNAs levels were independently validated and detected considerable variation in expression. All these results uphold the hypothesis that lncRNAs are tightly bound with the pluripotent state[25].

The importance of iPSC-enriched lncRNAs for iPSC derivation suggests a link between lncRNAs and the pluripotency network. They observed that lncRNAs appears to be controlled by pluripotency transcription factors in ESCs and iPSCs. In particular, after the knockdown of a lncRNA candidate (lincRNA regulator of reprogramming - lncRNA-RoR) the iPSC colonies decreased relatively to the control whereas progenitor cells were unaffected, showing that lncRNA-RoR is required for iPSC derivation[25].

Additionally, the authors investigated which cellular pathways were affected by lncRNA-RoR knockdown. The results suggested that the absence of lncRNA-RoR would led to a upregulation of genes involved in the p53 response, the response to oxidative stress and DNA damage-inducing agents, as well as cell death pathways. A knockdown of p53 partially rescued the apoptotic phenotype caused by ablation of lncRNA-RoR. These results led to the conclusion that lncRNA-RoR play an important role in promoting survival in iPSCs and ESCs, preventing the activation of cellular stress pathways including the p53 response[25].

1.4.1.1. ZEB2 and his natural antisense transcript lncRNA ZEB2NAT

An lncRNA that is pursued in our current laboratory is Zeb2NAT, an antisense of the Zeb2 coding gene.

The change of cells morphology is one of the first signs to be observed during cellular reprogramming. Epithelial-mesenchymal transition (EMT) is a cellular process that occurs in early development and cancer through down-regulation of epithelial genes, such as E-cadherin, and up-regulation of mesenchymal genes, such as Snail1 (zinc finger transcription factors homologue of Snail) [26]. During EMT epithelial cells lose their cell-cell adhesion and cell polarity and acquire invasive and migratory properties becoming mesenchymal cells. The mesenchymal-epithelial transition (MET), the reverse process of EMT, is also possible being crucial in reprogramming of fibroblasts to iPSCs. The reprogramming process is controlled by the balance between EMT and MET [27,28].

The expression of Zeb2 and Zeb1 transcription factor detected in mesenchymal cells is induced by Snail1 during EMT. The expression of Zeb2 assists the down-regulation of E-cadherin. Although Zeb2 mRNA levels remain stable, it is alternatively processed after Snail1 induction[29]. It's also known that in MET the expression of Zeb2 mRNAs is repressed by miR-200 family members[30].

When Snail1 is induced in cells, an intron located in the 5' -UTR of Zeb2 stop being processed. This intro contains an internal ribosome entry site (IRES) sequence crucial for the Zeb2 functional expression[29].

The conservation of this intron is controlled by the expression of a noncoding transcript that occurs in an antisense direction (Zeb2NAT) overlapping the 5' -UTR of Zeb2. Overexpression of Zeb2NAT in epithelial cells changes cell identity through increased Zeb2 protein levels (Figure 1.3)[29].

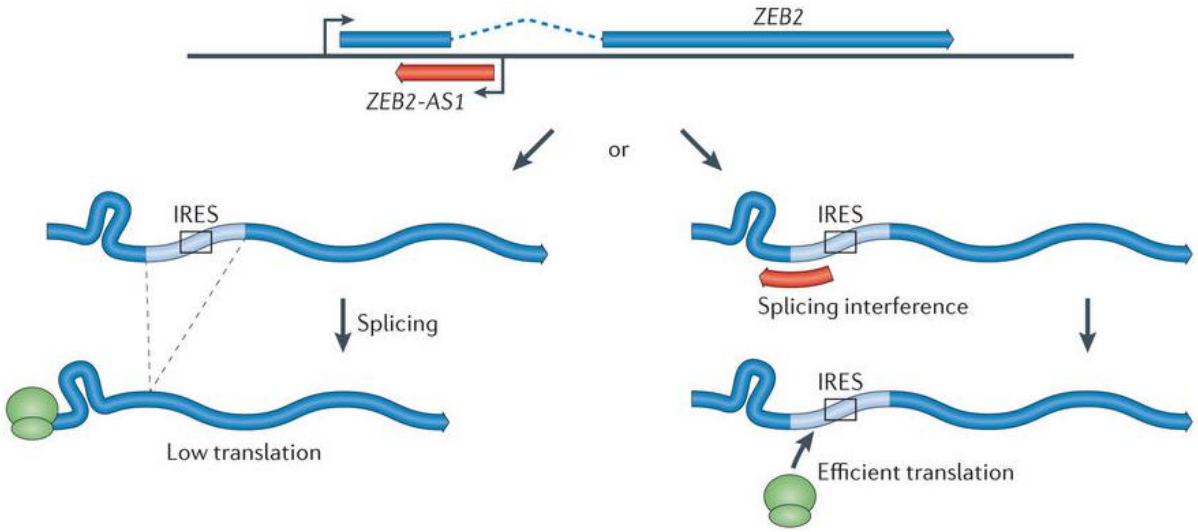


Figure 1.3 - Illustration of the regulation of Zeb2 through lncRNA Zeb2NAT. LncRNA Zeb2NAT is the antisense of the sense Zeb2 and has an important role in incrising the Zeb2 proteins levels by binding to the pre-Zeb2 mRNA, preventing the splicing of the intro that contain the IRES (Internal Ribosome Entry Site), that is essential for Zeb2 mRNA translation. The increase of Zeb2levels promotes the inhibition of E-Cadherin leading cells to an EMT[29].

Zeb2NAT increases the Zeb2 proteins levels by binding to the pre-Zeb2 mRNA, preventing the splicing of the intron that contains the IRES that is essential for Zeb2 mRNA translation. Zeb2 promotes the inhibition of E-Cadherin leading cells to an EMT (Figure 1.4). The knockdown of Zeb2NAT can lead the cells to a MET through the E-Cadherin expression that is increased with the decreasing of Zeb2 protein[30].

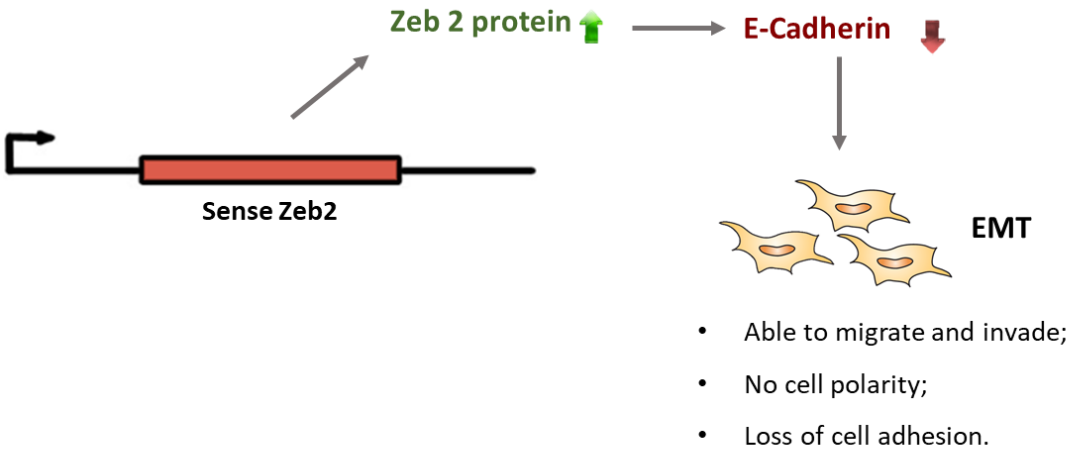


Figure 1.4 - Representation of the lncRNA zeb2 regulation of the epithelial-mesenchymal transition of cells through the E-Cadherin inhibition.

The inhibition of protein translation, such as Zeb2, can be performed by using a locked nucleic acid (LNA), that prevent mRNA-protein interactions[31]. Locked nucleic acids (LNA) are a class of RNA analogs in which the ribose ring is "locked" into a rigid bicyclic formation by a methylene bridge between 2' oxygen and the 4' carbon, conferring an enhanced performance and an increased breadth of applications[32].

LNAs can be used to overcome the difficulties of studying very short sequences and has greatly improved, and in many cases enabled, specific analysis and sensitive detection of DNA, micro RNA, ncRNA and other small RNA molecules. LNAs can be used to perform in situ hybridization, northern blotting, microarray analysis, real-time PCR, isolation, RNAi, gene repair/exon skipping, splice variant detection, antisense inhibition and more[32].

Due to the high affinity of LNA to complementary nucleic acids, LNAs: RNA duplexes are much more stable that those formed between DNA and RNA, makes LNAs as an extremely potent antisense inhibitor, both for in vitro and in vivo use[32]. This means that LNAs are one of the best options to perform the downregulation of lncRNAs[32].

1.4.1.2. LncRNA NORAD

Recently, Mendell and colleagues uncovered a novel lncRNA regulated by DNA damage, termed non-coding RNA activated by DNA damage (NORAD). NORAD plays a key role in maintaining genome integrity by modulating the activity of the RNA binding proteins PUM2 and PUM1 after DNA damage[33,34]. Mendell group was initially interested in identifying human lncRNAs that regulate the DNA damage response. To do so, they examined a set of previously identified mouse lncRNAs that are induced after doxorubicin treatment in a p53-dependent manner. Among all transcripts, they observed a poorly characterized 4.9kilobase (kb) unspliced murine lncRNA, annotated as

2900097C17Rik, that exhibits a high degree of evolutionary conservation in mammals and an unusually high abundance (500-1000 copies per cell). Human genome has a clear ortholog of this transcript, with 65% nucleotide identity to 2900097C17Rik, annotated in RefSeq as LINC00657 with 5.3kb. LINC00657 is a cytoplasmic lncRNA abundantly expressed in human cells and tissues[34].

They observed that the presence of NORAD in cells is fundamental to maintain genomic stability by sequestering PUMILIO proteins. Otherwise, the absence of NORAD would trigger chromosomal instability by repressing mitotic, DNA repair and DNA replication factors through the release of PUMILIO proteins[34].

NORAD affects genomic stability through its direct interaction with PUMILIO 2 (PUM2) and possibly PUMILIO 1 (PUM1), two human and mouse RNA-binding proteins belonging to the deeply conserved Pumilio-Fem3-binding factor (PUF) family. PUF proteins and more specifically PUM1 and PUM2 are capable of binding with high specificity to sequences in the 3' UTRs of target mRNAs through an 8-nt specific sequence (UGUANAUA), referred to as the PUMILIO response element (PRE), causing genomic instability[34](Figure 1.5).

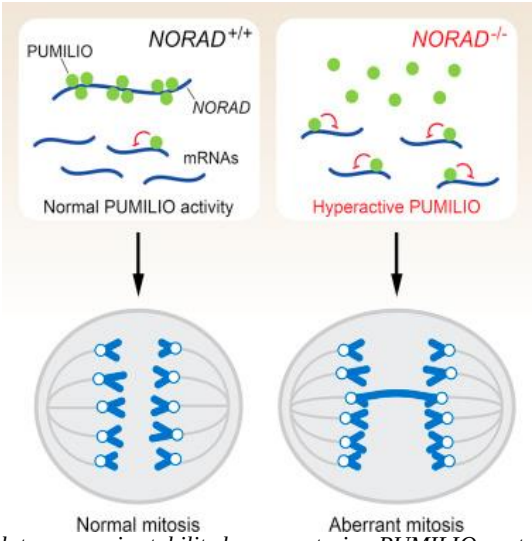


Figure 1.5 - LncRNA NORAD regulates genomic stability by sequestering PUMILIO proteins[34]

The presence of 15 PUMILIO-binding motifs, mostly located in five repeated 400-nt domains named NORAD Domains (ND1 through ND5), allows lncRNA NORAD to act as a potent molecular decoy for PUMILIO proteins (PUM1/PUM2), preventing the RNA-binding proteins from destabilizing their targets. In fact, NORAD is capable to bind approximately 7500-15000 PUMILIO molecules per cell, which represents 50%-100% of the total PUMILIO protein pool in the human colon cell line (HCT116)[34].

Recent findings from Mendell group revealed a new pathway to marked chromosomal instability (CIN), a phenotype characterized by the frequent gain or loss of chromosomes during mitosis. Mendel and his lab findings support what they previously suggest earlier, that PUMILIO are an important regulator of genomic stability acting as repressors of a set of genes whose expression is necessary to maintain chromosomal stability, such as DNA replication and repair as well as key mitotic factors genes. They hypothesized that the deregulation of the targets under conditions of PUMILIO hype- or hypoactivity could induce a state of severe genomic instability as observed upon NORAD loss function, PUM1/2 overexpression or PUM1/2 inactivation[34].

Another study, now from Pera group, identified an intriguing expression of PUM2 protein in embryonic stem cells. In fact, they observed that PUM2 can form a stable complex with DAZ protein and its homolog, DAZL (DAZ-Like) protein in diverse organism, including humans. DAZ/DAZL are proteins necessary for germ cell development in males and females (DAZ protein is only expressed in males)[35].

They observed that the capacity of PUM2 protein to form a stable complex with DZA through RNA binding and protein-protein interaction makes PUM2 a requirement to maintain germ line stem cells in *Drosophila* and *Caenorhabditis elegans*. When the PUF repeat 8 of PUM2 was deleted, this interaction between PUM2 and DAZ decreased suggesting a link between both proteins, supporting the first observation of PUM2 importance in germ line maintenance. They also observed an interchange between DAZ and DAZL homologs when a human DAZ transgene was able to partially rescue a mouse DAZ mutation[35].

This findings indicates that PUM2 function can pass by maintaining the human germ cell lineage giving rise to mature germ cells in men and women as PUM2 is expressed throughout the development of the female and male germ cell lines. They hypothesized that PUM2 together with DAZ and DAZL proteins can act as a translational regulator in the germ cell lineage[35].

However, the PUM2 role in the germ cells remains unclear suggesting the need for a further investigation to uncover the PUMILIO proteins functions and how are regulated in stem cells[34,35].

1.5. Objectives

The main objective of this project passed by understanding the cellular reprogramming of human aged cells. Previously results from our laboratory reported a decreased reprogramming efficiency when reprogramming adult somatic cells from mice. Here, we recently characterized an lncRNA capable of enhancing reprogramming of aged cells. When this lncRNA, ZEB2NAT, was down-regulated with LNAs we could achieve higher reprogramming efficiencies. Unfortunately, how aging influences the reprogramming of human cells and how Zeb2NAT works on human cell reprogramming is still unknown.

Actually, an objective of this project passed by testing if aging is a reprogramming barrier in human cells and if it's possible to enhance reprogramming by downregulating Zeb2 and/or his antisense transcript ZEB2NAT.

Furthermore, was also intended to understand how a novel lncRNA, called NORAD, could affect stem cells properties, viability and also chromosomal stability since previous results have already shown that PUM2 have an important role in maintain germ cell lineage. Therefore, a last objective of this project was also to verify if there NORAD can regulate the stem cell stability and improve stem cells proprieties through PUM2 protein regulation.

2. Materials and Methods

2.1. Cell lines

Human Fibroblasts

The reprogramming of fibroblasts into hiPSCs was performed using two human fibroblasts cell lines, the human embryonic fibroblasts (WI38) and the 3yr human fibroblasts (3yr, supplied by ATCC), both with low passage (passage 2 and 4) and high passage (passage 7).

Human fibroblasts were maintained in Dulbecco's Modified Eagle Medium (DMEM, Gibco™) supplemented with 10% of Fetal Bovine Serum (FBS), 1% of L-Glutamina and 1% of Penicillin Streptomycin (Pen/Strep) under sterile conditions and incubated at 37°C and 5% CO2.

Fibroblasts were passed when reached a 90/95% of confluence by washing with Phosphate Buffered Saline (PBS) and detached with trypsin. Posteriorly, cells were resuspended in DMEM medium to inhibit trypsin and plated. When needed, fibroblasts were freeze by centrifuging the cells resuspended in DMEM medium at 1250 rpm during 5 minutes at room temperature (RT). The supernatant was removed and the pellet was resuspended in 1ml of 90% FBS + 10% DMSO (Sigma) and added to a cryogenic vial. Cells were stored at -80°C during 2 or more days and then stored at liquid nitrogen. The concentration of cells used per vial was between 1×10^6 and 1×10^7.

To perform a cellular reprogramming assay, human fibroblasts were counted during their passaging by mixing 20µl of a trypan blue solution with 20µl of cells resuspended in DMEM medium, before plating. The solution was homogenized and loaded into a Neubauer Chamber. As the ratio between cells and trypan blue used was 1:1, the formula to calculate cell concentration used was $mean \times 2 \times 10^4$, were the mean is calculated by the sum of the farthest quadrants of the extremities and divided by 4.

If needed to thaw fibroblasts to perform an assay, the frozen cryogenic vial would be withdrawn from the liquid nitrogen and thawed at 37°C for 3 minutes at maximum. All the volume of cryogenic vial was added to a falcon with PBS and resuspended. A pellet of cells was formed by centrifuging the falcon at 1000-1250 rpm for 5 minutes. Afterwards the supernatant was aspirated and the pellet was resuspended in DMEM medium and plated.

Human embryonic kidney 293T cell line (293T)

The 293T cell line (ATCC) was used to produce the viral medium to perform cellular reprogramming experiment. 293T cell line were maintained in DMEM supplemented with 10% of FBS, 1% of L-Glutamina and 1% of Penicillin Streptomycin under sterile conditions and incubated at 37°C and 5% CO2.

293T cells were maintained, passed, freeze and thawed by using the exact procedure as used for human fibroblasts. However, pre-coated culture petri dishes with gelatin were needed for 293T survival. Coating culture petri dishes was performed by adding gelatin 0.1% and incubated at 37°C during 10-15 minutes. After removing the gelatin, fresh DMEM cell medium was added along with 293T.

Human induced pluripotent stem cells (hiPSCs)

The human iPSCs obtained through cellular reprogramming of the human fibroblasts mentioned above were grown in mTeSR™1 medium (Gibco™) containing 90% of mTeSR™1 Basal medium, 9,5% of mTeSR™1 5X supplement and 0.5% of Pen/Strep, under sterile conditions and incubated at 37°C and 5% CO2. To maintain hiPSCs in culture, the medium was changed every day and hiPSCs were plated in pre-coated culture petri dishes with Matrigel™ (Corning).

Before hiPSCs clones started to touch each other, hiPSCs were passed by using one of the two protocols: hiPSCs passage with EDTA and hiPSCS passage with trypsin.

In the hiPSCs passage with EDTA protocol, cells were washed with PBS first. PBS-EDTA was added after removing PBS and incubated at RT for 3 minutes. Cells were rinsed and scraped from bottom, with care to avoid to getting single cell suspension. Cells were transferred to a 15ml falcon tube with additional medium according to the dilution wanted and resuspended carefully. Cells were plated in pre-coated culture petri dishes with Matrigel™. If it were necessary to improve the recovery and growth of hiPSCs due to a single suspension a 1µl of ROCK inhibitor (Axon) per 1ml of mTeSR™1 medium would be added. Rock inhibitor or selective Rho-associated kinase inhibitor increase the dissociated human stem cells survival by inhibiting apoptosis [36].

In the hiPSCs passage with Trypsin protocol, cells were washed with PBS. After removing PBS, cells were incubated with trypsin for 2-3 minutes at 37°C. Afterwards a few drops of FBS serum were added to inhibit trypsin effect. Cells were resuspended and centrifuged at 1250 rpm for 5 minutes. Supernatant was removed and the pellet resuspended in mTeSR™1 medium. Afterwards hiPSCs were plated in a plate previously coated with Matrigel™. To improve the recovery and growth of hiPSCs due to single cell suspension a 1µl of ROCK inhibitor per 1ml of mTeSR™1 medium would be added.

Culture petri dishes were previously pre-coated with Matrigel™ by thawing an aliquot of Matrigel™ on ice. In order to dilute Matrigel™ aliquot, DMEM-F12 medium (Gibco™) was added. The diluted Matrigel™ was added to culture petri dishes and incubated at 37°C for 20-30 minutes. Matrigel™ was then aspirated and mTeSR™1 was added.

When needed to freeze the hiPSCs resuspension from hiPSCs passage was centrifuged at 1250 rpm for 5 minutes. The supernatant was removed and the pellet resuspended in 90% KnockOut™ Serum (KSR) + 10% DMSO or 90% FBS (stem cell ready) + 10% DMSO solution. Cells were stored at -80°C during 2 or more days and then stored at liquid nitrogen.

The PBS-EDTA (0.5mM) used was made by diluting 200µl of 0.5 M EDTA with 200ml of PBS and the osmolarity was adjusted by adding 0.36g of NaCl, and sterilized using a filter-sterile. The freezing medium used contained KSR with 10% of DMSO.

Mouse stem cells E14

To study NORAD role in stem cells a mouse stem cell E14 cell line (a kind gift from Dr. Manuel Serrano, CNIO, Madrid) was used. E14 was maintained, under sterile conditions and incubated at 37°C and 5% CO2, in KnockOut™ DMEM medium with Leukemia inhibitor factor (KSR+LIF). KSR medium was made by supplementing basal DMEM medium with 15% of KSR (Gibco™), 1% of L-Glutamina, 1% of Pen/Strep, 1.2% of non-essential amino acids solution (NEAA) and 0.18% of Beta-

mercaptoethanol. For an aliquot of 50ml of KnockOut™ DMEM medium was also added 0.012% of leukemia inhibitor factor supplement (LIF, Millipore) (KSR+LIF).

E14 were passed by washing with PBS. PBS was removed and the cells were trypsinized by adding a few drops of trypsin and incubated for 3-5 minutes at 37°C. Posteriorly, cells were resuspended in PBS or FBS (stem cell ready) and centrifuged at 1250 rpm during 5 minutes. Supernatant was removed and the pellet was resuspended in KSR+LIF medium and plated. Culture petri dishes were pre-coated with gelatin using the same coating gelatin protocol for 293T. When needed, the E14 were frozen by centrifuging at 1250 rpm during 5 minutes the E14 resuspension in PBS or FBS (stem cell ready). The supernatant was removed and the pellet resuspended in 90% KSR + 10% DMSO or 90% FBS (stem cell ready) + 10% DMSO solution. Cells were stored at -80°C during 2 or more days and then stored at liquid nitrogen.

To thaw E14, the frozen cryogenic vial would be withdrawn from the liquid nitrogen and thawed at RT. All the volume of cryogenic vial was added to a falcon with PBS and resuspended. Afterwards the supernatant was aspirated and the pellet was resuspended in KSR+LIF medium and plated in pre-coated plates with gelatin.

2.2. Cellular Reprogramming

293T transfection (viral medium production)

Cellular reprograming of human fibroblasts was performed based on the protocol of Li group [37]. 293T cells were seeded with a confluence of 2×10^6 on a 100 mm culture dish previously coated with gelatin and incubated 24 hours at 37°C. After the 24 hours of incubation two mixes were made:

- A containing lentiviral transfer vector and second generation packaging vectors;
- B with 600µl of DMEM and 50µl of X-tremeGENE™ 9 (Roche).

Both mixes were mixed together and incubated at RT for 20 minutes after a gently vortex. Transfection mixture was added dropwise to cells and incubated overnight at 37°C. 24 hours after, DMEM medium was replaced by new DMEM medium. Viral medium was collected 36h - 48 hours after transfection for first infection and new fresh medium was added to 293T. After 24 hours a second round of viral medium was collected to perform the second infection.

For viral medium production was used 10µg of transfer vector. FUW-OSKM plasmid (Addgene #20328), pkp332 plasmid (Addgene #21627) and a positive infection control, the green fluorescence protein plasmid (provided by Edgar Gomes lab of IMM). For the packaging vectors was used a second generation packaging system composed by 5µg of both pMD2.G (Addgene #12259) and psPAX2 (Addgene #12260) plasmids. All plasmids were produced using STBL3 competent bacteria (Thermo Fischer Scientific), following supplier protocol.

Fibroblasts Infection (transduction of transcription factors)

The target cells were plated at least 24 hours prior to viral infection with approximately 90% of confluence. Viral medium collected was filtrated by passing through a 0.45µM low protein-binding

filter and mixed with 1:1000 polybrene (Sigma-Aldrich). Cells were infected with the first viral medium collected and incubated for 24 hours at 37°C. The next day a second infection was performed in cells with the second viral medium collected and incubated for 24 hours at 37°C. Afterwards viral medium was removed and cells were washed with PBS before fresh DMEM medium was added.

2.2.1. Downregulation of Zeb2 and Zeb2NAT

LNAs GapmeRs transfection

Cells were plated in a 12 well plate at least 24 hours prior to LNAs GapmeRs transfection. Two mixes were made: mix A containing 100µl of Opti-MEM (Gibco™) with 1µl of specific LNA (25µM) and mix B containing with 100µl of Opti-MEM and 2µl of Lipofectamine® RNAiMAX Reagent (Thermo Fisher Scientific). Both mixes were incubated for 5minutes at RT. Mix A was gently transferred to mix B and incubated for 20 minutes at RT. Mixture of both mixes was added to the cells and incubated for 24 hours at 37°C. A second transfection was performed in the next day following the same steps as the day before. Afterwards viral medium was removed and cells were washed with PBS before fresh DMEM medium was added.

Specific LNA-GapmeRs (EXIQON™) used: Zeb2, Zeb2NAT and a non-specific control (scramble) as negative control.

MZeb2_12(h/m): 5'- GTTAGCCTGAGAGGAG-3';

Zeb2NAT_1(h): 5'- TAATTTACTTAGAGAC-3';

Zeb2NAT_18(h): 5'- GTCCAGAAATTCATC-3';

MZeb2NAT_12(m/h): 5'-TTAGTGATGAGGATA -3';

Scramble: 5'-AACACGTCTATACGC -3'.

2.3. Molecular Biological Techniques

Real-time quantitative PCR (RT-qPCR)

RNA extraction was performed by using pureZOL™ (Bio-Rad), following the standard instructions from manufacturer. The RNA obtained was treated with DNase kit (Roche), to eliminate any DNA that could contaminate the samples, followed by a reverse transcription to transcribe the total RNA into complementary DNA (cDNA), following manufacturer instructions (Roche).

RT-qPCR was performed using a ViiA7 of 96 wells and 384 wells equipment from Applied Biosystems™. This technique was used for gene expression analysis of Zeb2, Zeb2NAT and hOct4 and hSox2. This technique consists in quantifying the amount of product DNA that it is being amplified through the inclusion of a fluorescence reporter molecule in each reaction. For this fluorescence reporter molecule was used iTaq™ Universal SYBR® Green Supermix (Bio-Rad). ViiA7 records the fluorescence values emitted by SYBR Green during each cycle of the amplification

process. The result recorded is given in threshold cycles (Ct) that corresponds to the first statistically significant fluorescence detection which is above the baseline or background. This is automatically defined by the software (RT-qPCR ViiA7 –Applied Biosystems). All the cycles temperatures was defined by the ViiA7, already predefined by manufacturer and the data was analyzed using the $2^{(-\Delta\Delta Ct)}$ method[38].

Housekeeping genes sequences used:

hbeta-actin-F: 5' -TGACGTGGACATCCGCAAAG -3';

hbeta-actin-R: 5' -CTGGAAGGTGGACAGCGAGG -3';

hGAPDH-F: 5'-GACAGTCAGCCGCATCTTCT -3';

hGAPDH-R: 5'-TTAAAAGCAGCCCTGGTGAC -3'.

Mastermix: iTaq™ Universal SYBR® Green Supermix + Primer F + Primer R + water (RNase/DNAse free).

Primers sequences used:

hZEB2(pair1)-F: 5'-TGTTTCTGCAAGTGCCATCC-3';

hZEB2(pair1)-R: 5'-ACACTGAAGCTGGTGCAAAG-3';

hZEB2(pair2)-F: 5'-TCCACCTCAAAGCGCATTTC-3';

hZEB2(pair2)-R: 5'-AGCAAAAGTAGCTGCTCCAG-3';

hNAT(pair1)-F: 5'-ATAATTGAAGCGCCCTGAGC-3';

hNAT(pair1)-R: 5'-GCACACACCCTAATACACATGC-3';

hSox2-F: 5'-CAAGATGCACAACTCGGAGATCAG-3';

hSox2-R: 5'-GGGCAGCGTGTACTTATCCTTCTTC-3';

hOct4-F: 5'-GAGGATCACCCTGGGATATACA-3';

hOct4-R: 5'-CGATACTGGTTCGCTTTCTCTT-3'.

Flow Cytometry

Flow cytometry was performed using a BD Accuri C6 flow cytometer. Cells were prepared following the standard protocol given by the manufacturer (BD Biosciences™). Flow cytometry was used to quantify the amount of cells expressing GFP (by calculating absolute cell concentration (per unit sample volume) automatically). For GFP detection was used a Blue laser (488nm) with a 530/30 filter. The data was analyzed by using FlowJo® software.

Alkaline Phosphatase Detection (AP staining)

AP staining was performed using Alkaline Phosphatase kit (Millipore). Cells were prepared according to the instructions from the manufacturer. AP kit was used to detect high level of alkaline phosphatase activity in human induced pluripotent stem cells through a fluorometric assay. A red or purple color appears when in presence of iPSCs otherwise cells appear colorless.

Immunofluorescence

Reprogrammed hiPSCs were picked and cultured in Matrigel™-coated cover slips. After reaching a good confluence hiPSCs were fixed with paraformaldehyde (3.7%) for 10 minutes at RT and then permeabilized with Triton-X (0.2%) for 5 minutes and blocked with BSA (2%) for 45 minutes at RT. For hiPSCs detection was used a human antibody against Nanog (Thermo Fisher Scientific) and a human antibody against Sox2 (R&D Systems). For the nuclear cell stain was used DAPI (Sigma-Aldrich). Fluorescence microscope images were acquired using a Zeiss Observer microscope.

Teratoma formation assay

hiPSCs were expanded in 100 mm cell culture plates pre-coated with Matrigel™ until a good confluence was reached. Cells were collected by PBS-EDTA and resuspended in 600µl of Matrigel™. Afterwards hiPSCs were injected subcutaneously into dorsal flanks of NSG mice (IMM rodent facility). Mice were sacrificed with anesthetic overdose for teratoma isolation and the teratomas were analyzed in the histology facility of IMM.

NORAD plasmid for transfection construction

Two pairs of oligonucleotides for NORAD deletion were specifically designed with a mouse homology sequence from a mice 4.9kb lncRNA NORAD (290097C17rik).

The two pairs of oligonucleotides were designed with the following sequence:

Oligo1NORAD-F: 5’-CTTCGTGAACCAGCACGTAGCACAA-3’;

Oligo1NORAD-R: 5’-AAACTTGTGCTACGTGCTGGTTCAC-3’;

Oligo2NORAD-F: 5’-CTTCGCCCCTCGGGCTCACCCGGCG-3’;

Oligo2NORAD-R: 5’AAACCGCCGGGTGAGCCCGAGGGGC-3’.

Each oligo was phosphorylated before annealing using a T4 polynucleotide kinase (NEB). Was used the Px459 vector to insert each annealed pair. Firstly, Px459 vector was cut with BbsI enzyme. An electrophoresis on 1x TAE with 0.8% agarose gel stained with Midori Green Advance (Nippon Genetics Europe) was performed to confirm if Px459 vector was cut correctly. The corresponding vector DNA was extracted from gel and purified using NZYGelpre kit (NZYTech). After purification,

linearized vector Px459 was de-phosphorylated using rapid alkaline phosphatase, inactivated at 75°C for 5 minutes. With both pairs of annealing oligos and Px459 vector prepared, a ligation of both with T4 DNA ligase was performed and incubated at 25°C for 1 hour.

After, ligation, competent bacteria DH5α was transformed with the insert and plated in Lysogeny broth (LB) + agar previously prepared with 0.1% of ampicillin for clone selection and incubated at 37°C, overnight. Clones were picked and a GeneJET Plasmid Mini-prep (Roche) was performed. DNA samples were quantified using NanoDrop 2000 (Thermo Fisher Scientific). Extracted Plasmids were then digested with AGE1 and BbsI restriction enzyme. An electrophoresis on 1% TAE with 0.8% agarose gel stained with Midori Green Advance was performed. If the insert was incorporated in the vector, only one band of 9kbs would be visible, if not, two bands of 8kbs and 1kbs would appear. After positive confirmation of the insert incorporation, plasmids were sent for Sanger sequencing.

After confirming the correct sequence, the construct was cloned by transforming the competent bacteria DH5α with the insert. Clones were picked in the following day and a GeneJET plasmid Midi-prep (Roche) was performed. DNA samples were quantified using NanoDrop 2000. All reagents and kits were used following the manufactures protocols.

NORAD transfection

E14 were plated in a 6 well plate at least 24 hours prior to transfection. Two mixes were made:

- Mix A containing 100µl of Opti-MEM (Gibco^{TM}) with 1µl of NORAD plasmid (25µM)M;
- Mix B containing with 100µl of Opti-MEM and 2µl of Lipofectamine® RNAiMAX Reagent.

Both mixes were incubated for 5 minutes at RT. Mix A was gently transferred to mix B and incubated for 20 minutes at RT. Mixture of both mixes was added to the cells and incubated for 24 hours at 37°C. A second transfection was performed in the next day following the same steps as the day before. Afterwards viral medium was removed and cells were washed with PBS before fresh KSR+LIF medium was added.

After both transfections were performed, E14 cells were maintained in KSR+LIF. Two days after, 1µg of Puromycin antibiotic was added to KSR+LIF medium for clone selection. Only the clones with the NORAD plasmid incorporated would survive, while the E14 without NORAD plasmid would die. Cells were kept with KSR+LIF medium with Puromycin for only 24 hours. Afterwards, the clones that survived were picked and expanded. After reaching a good confluence, a genomic DNA extraction was performed by using a cell lysis buffer composed by 0.2µg/µL of proteinase k (Sigma-Aldrich), 100mM NaCl, 10mM Tris pH 8.0, 25mM EDTA pH 8.0 and 0.5% SDS. Afterwards a PCR was performed to verify if the deletion was accomplished using Taq NEB enzyme, a melting temperature of 65°C and the following two pairs of NORAD primers sequences:

NORAD(pair1)-F: 5'-ACCTGGAAAGGCAAACACTG-3';

Primer1(pair1)-R: 5'-ATGCTTCTGCAGGCTTGAAT-3';

Primer2(pair2)-F: 5'-ACCTGGAAAGGCAAACACTG-3';

Primer2(pair2)-R: 5'-TATTCTCCCCTTGGCCTTCT-3'.

An electrophoresis on 0.5x TBE with 1% agarose gel stained with Midori Green Advance (Nippon Genetics Europe) was performed using the PCR product to confirm if the NORAD gene was deleted.

Electroporation

Electroporation was performed using an Electroporation kit (Thermo Fisher Scientific). Mouse E14 cells were prepared according to the standard protocol given by the manufacturer. Electroporation is a technique that creates temporary pores in cell membranes through electrical pulses. The electroporation performed to delete NORAD gene from E14 cells was used with the following conditions: 3 pulses of 10ms with a voltage of 1400V.

After electroporation was performed, the cells were added back to the plates pre-coated with gelatin. Two days after electroporation, medium was changed to KSR+LIF+Puromycin. Puromycin (Sigma-Aldrich) is an antibiotic used for clone selection. Only the cells with the plasmid incorporated survive, while the others (Specially the control without transfection of NORAD) die. Cells were kept with KSR+LIF+Puromycin for only 24 hours. Afterwards, the clones that survived were picked and expanded. After reached a good confluence, a genomic DNA extraction was performed. Afterwards a PCR was performed to verify if the deletion was accomplished using Taq NEB enzyme, a melting temperature of 65°C and the following two pairs of NORAD primers sequences:

An electrophoresis on 0.5x TBE with 1% agarose gel stained with Midori Green Advance was performed using the cDNA synthesized to confirm if the NORAD gene was deleted.

Genomic DNA extraction

E14 clones were picked with a pipette point 6 to 9 days after NORAD transfection for a positive confirmation that deletion of NORAD gene was accomplish with success. A genomic extraction of DNA was performed. Clones were lysate with lysis buffer with proteinase k (Sigma-Aldrich) and incubated overnight at 56°C. Genomic DNA was purified by adding 55µl of 3M Sodium Acetate pH 5.2, 500µl of phenol:chloroform:isoamyl alcohol (Thermo Fisher Scientific). Mix was vigorously mixed by shaking and centrifuged at 13000 rpm for 10 minutes at RT. Clear top phase was carefully transferred to a pre-labeled 1.5ml Eppendorf tube and 500µl of Chloroform (Merck) was added and mixed vigorously by shaking. Mix was centrifuged at 13000 rpm for 10 minutes at RT. Again, clear top phase was transferred carefully to a pre-labeled 1.5 ml Eppendorf tube. 450µl of Isopropanol was added and Eppendorf tube was inverted and swirled several times gently and incubated at -20°C, overnight. In the next day was centrifuged at 13000 rpm for 30 minutes at RT and supernatant was removed carefully without disrupting the pellet. Pellet was washed by adding 700µl of 70% ethanol and centrifuged at 13000 rpm for 5 minutes at RT. The pellet was centrifuged again at 13000 for 5 minutes at RT. Supernatant was removed carefully and Eppendorf was left open at RT until pellet was dry (during approximately 15-30 minutes). Afterwards, pellet was resuspended by adding 200µl of DNase/RNAse free dH_2O and left around 3 hours at 37°C until the pellet was well resuspended. _DNA was quantified using Nanodrop 2000 (Thermo Fischer Scientific).

3. Results and Discussion

To be able to attain our goals, optimization of the cellular reprogramming was a first requirement. To optimize the protocol, the experiments performed were tested with changes in different factors that could have an influence in the reprogramming efficiency, such as the concentration and passage of 293T cells, WI38 and 3yr human fibroblasts, the volume of the transfection reagent X-tremeGENE™ 9 (Roche) and the timing to perform each step of the protocol.

To perform the cellular reprogramming, fibroblasts were infected using a viral DMEM medium containing viral particles for the expression of the four Yamanaka factors. Two plasmids candidates to express transcriptional factors were tested, FUW-OSKM (catalog #20328, addgene) and pKP332 (catalog #21627, addgene) (Figure 3.1, A and B). The FUW-OSKM, referred here as OSKM, is a mouse lentiviral plasmid expressing mouse Oct4, Sox2, Klf4 and c-Myc while pKP332 is a human lentiviral plasmid expressing human Oct4, Sox2 and Klf4 for iPS cell generation. After testing both lentiviral plasmids in both human fibroblasts cell lines, only OSKM plasmid was able to induce the formation of hiPSCs clones (Figure 3.1, C). Due to this, OSKM plasmid was chosen to be used for all cellular reprogramming assays. However, even with the use of the OSKM lentiviral plasmid the iPSC clones generated didn't show an exact morphology of a true iPSC urging for the need of an optimization of reprogramming protocol.

Was also tested an infection using a second generation packaging system and a third generation packaging system, both based on the human immunodeficiency virus, HIV, [39,40], to verify which generation would be more efficient to reprogram. Each generation packaging system needs a lentiviral vector containing the specific insert, an envelope vector and while second generation packaging system needs only one packaging vector that contains all necessary viral structure proteins, the 3rd generation packaging system needs to have two packaging vectors. However, despite the complexity, the 3rd generation packaging system provides a maximal biosafety in the transfection. Although, the 3rd generation packaging system mix only supports 3rd generation lentiviral expression vectors unlike the 2nd generation that supports both 2nd and 3rd lentiviral expression vectors [41,42].

We observed that the second generation had a much better efficiency with the advantage of being a simplistic system, comparatively to the third generation. To perform the assays detailed hereafter only the second generation packaging system was used.

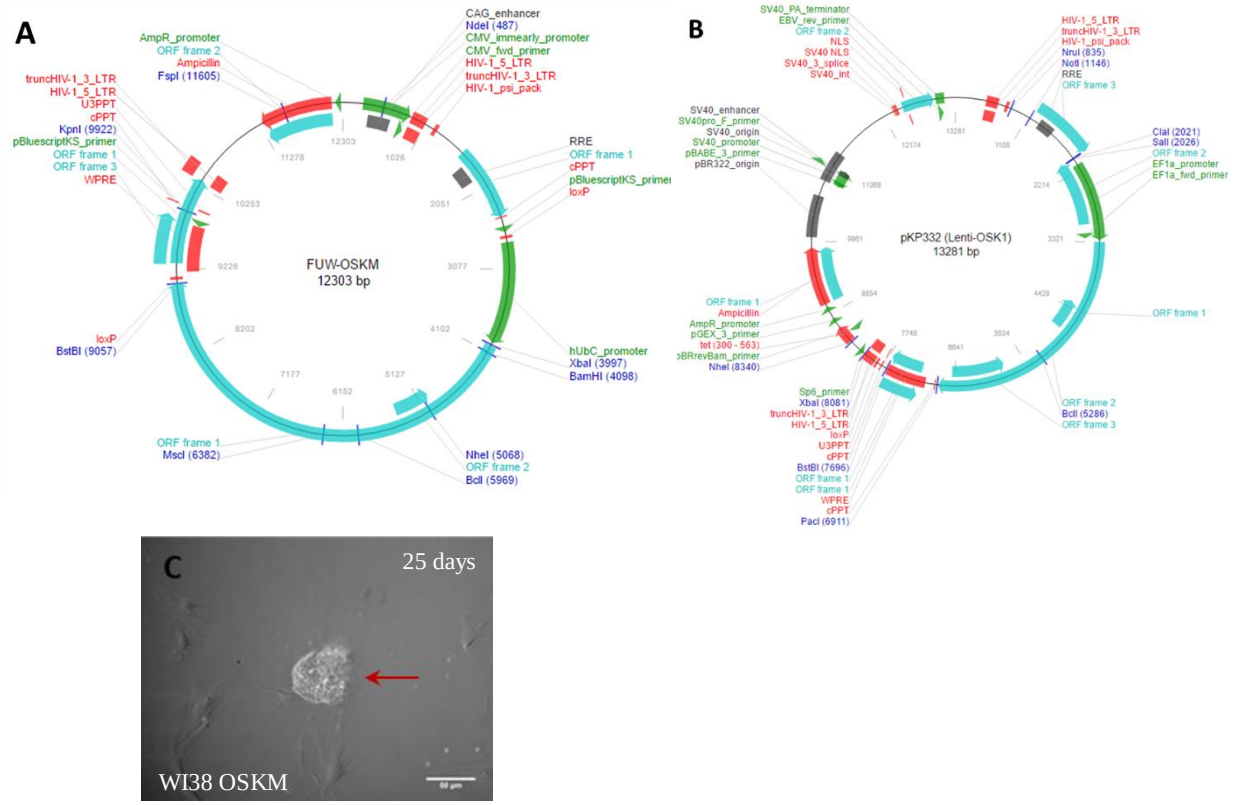


Figure 3.1 - Cellular reprogramming to test efficiency of OSKM and Pkp332 plasmids in reprogramming. (A) FUW-OSKM vector structure (B) pKP332 vector structure (C) WI38 Clone generated 25 days after transduction with OSKM plasmid. Amplification x10. Scale bar was set for 50µm. Image acquired through the Carl Zeiss PrimoVert microscope. Image was treated posteriorly with Fiji software.

A RT-qPCR was performed using the cells infected with viral medium containing OSKM particles to determine the levels of expression of the human Oct4 (hOct4) and human KLF4 (hKLF4), two of the four transcriptional factors. Samples were prepared by following the RT-qPCR protocol detailed in materials and methods. The RT-qPCR results obtained for hOct4 are presented in Figure 3.2. The hKLF4 results are presented in annex 1, Figure 6.1. Was performed the same procedure using the cells infected with viral medium produced using the pKP332 plasmid, also showing a high expression of hOct4 (Annex 1, Figure 6.2). Despite this, no hiPSCs was generated, probably due to the absence of c-MYC (one of the four transcriptional factors) conditioning the cellular reprogramming.

The results, from Figure 3.2, show an increasing of hOct4 and hKLF4 levels in both human fibroblasts comparatively with their controls. This means that transduction using OSKM plasmid was working efficiently. However, WI38 human fibroblasts showed a higher levels of hOCt4 and hKLF4, generating more hiPSCs clones comparatively with 3yr human fibroblasts.

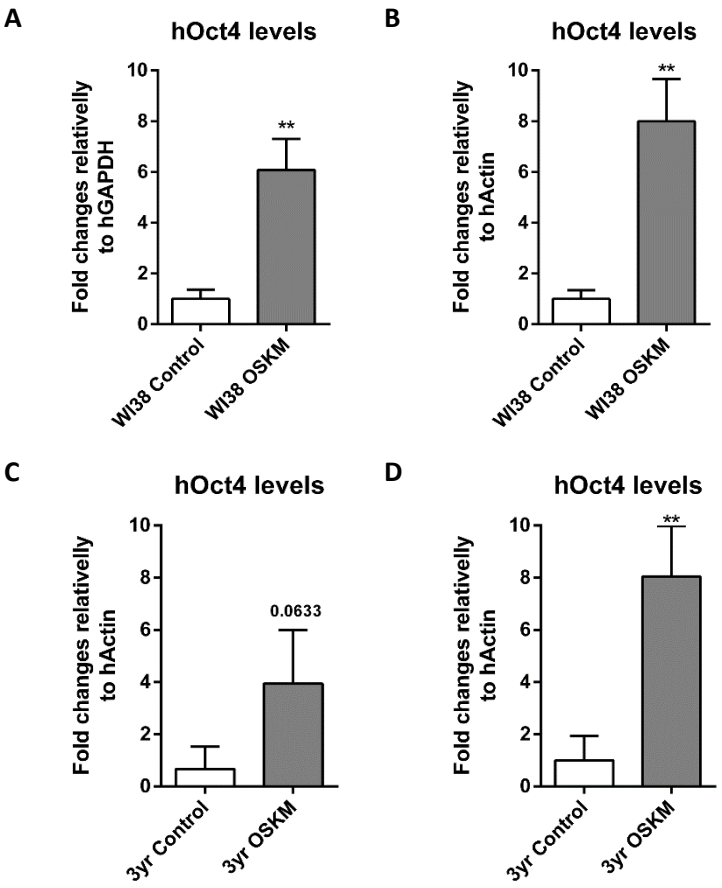


Figure 3.2 - RT-qPCR results of hOCt4 expression levels of WI38 and 3yr human fibroblasts (after transduction using OSKM plasmid) presented as $\Delta\Delta C_t$ normalized using non OSKM lentiviral transduced cells as control and hGAPDH and hActin as housekeeping gene. (A and B) WI38 human fibroblasts hOct4 levels using housekeeping hGAPDH (A) and hActin (B). (C and D) 3yr human fibroblasts hOct4 levels using housekeeping hGAPDH (C) and hActin (D). P-value ≤ 0.05; p-value ≤ 0.01**; p-value ≤ 0.001***.*

To determine the infection efficiency of lentiviral plasmids for each cellular reprogramming experiment performed, was used a lentivirus for the expression of a green fluorescence protein (GFP) as the positive control for the infection. Cells infected with GFP emitted a signal between a wavelength of 500-550nm in the visible spectrum. The infection efficiency was obtained through flow cytometry analysis, which quantifies the amount of cells alive that were expressing GFP. Cells were prepared by using flow cytometry protocol. The flow cytometry results referred to GFP expression of cells from Figure 3.3 are shown in Figure 3.4.

Figure 3.3 show a GFP infection of both human fibroblasts from a cellular experiment performed. Both cell lines expressed green fluorescence, however, the GFP signal in the 3yr old cell line was weaker.

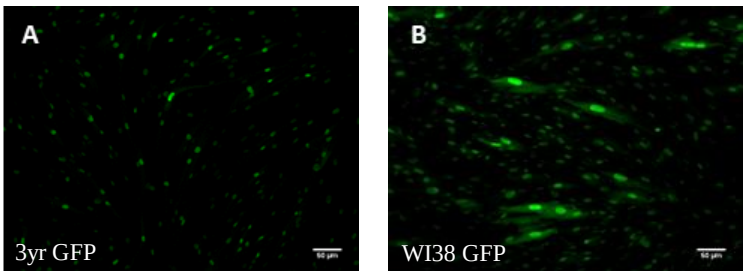
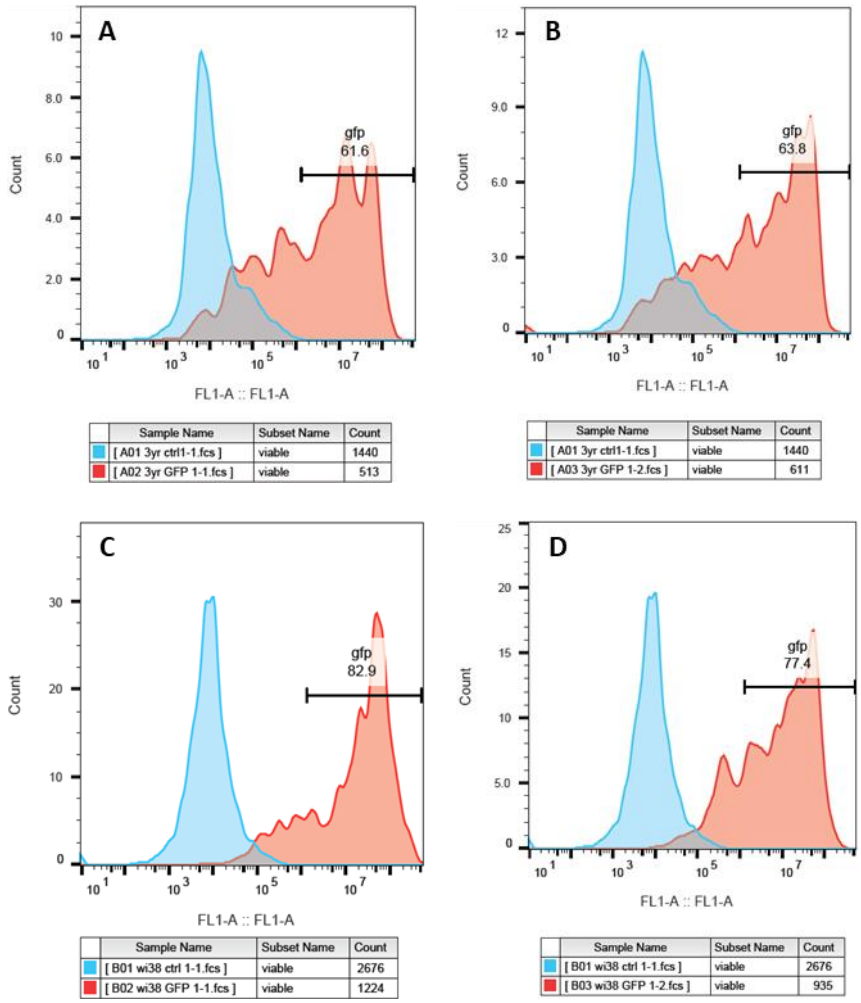


Figure 3.3 - GFP expression after a cellular reprogramming experiment with both 3yr (A) and WI38 (B) human fibroblasts performed using cellular reprogramming protocol. Scale bar was set for 50µm. Images acquired through the Zeiss Cell Observer fluorescence microscope, amplification x4. Images were treated posteriorly by using Fiji software.

The preliminary results presented in Figure 3.4 show a high GFP infection efficiency for both cell lines. However, younger human fibroblasts WI38 had a higher percentage of cell infected (mean= 80.2%) comparatively to the older human fibroblasts 3yr (mean=62.7%), correlating with the first microscope observation.



Sample Name	Subset Name	Count
[A01 3yr ctrl1-1.fcs]	viable	1440
[A02 3yr GFP 1-1.fcs]	viable	513

Sample Name	Subset Name	Count
[A01 3yr ctrl1-1.fcs]	viable	1440
[A03 3yr GFP 1-2.fcs]	viable	611

Sample Name	Subset Name	Count
[B01 wi38 ctrl 1-1.fcs]	viable	2676
[B02 wi38 GFP 1-1.fcs]	viable	1224

Sample Name	Subset Name	Count
[B01 wi38 ctrl 1-1.fcs]	viable	2676
[B03 wi38 GFP 1-2.fcs]	viable	935

Figure 3.4 - Flow cytometry histogram results of 3yr and WI38 human fibroblasts expressing green fluorescense plotting in red. 3yr and WI38 control plotted in blue.. (A) and (B) 61.6% and 63.8% of viable 3yr human fibroblasts expressing GFP, correspondently; mean= 62.7%; (C) and (D) 82.9% and 77.4% of viable WI38 human fibroblasts expressing GFP, correspondently, mean= 80.2%. Analysis performed using BD Accuri C6 with a cell count of 3000.

3.1. Optimization of the Cellular reprogramming protocol

The optimization of cellular reprogramming was the first requirement to be made in order to test the mentioned objectives. This was achieved by changing a series of factors that could be linked to the reprogramming efficiency. This was extremely necessary due to the fact that the clones generated from cellular reprogramming with OSKM plasmid were small and impossible to expand. The clones weren't dividing, indicating the absence of true hiPSCs features. More probably, the clones were in some transition from human fibroblasts to hiPSCs.

Reprogramming human fibroblasts into hiPSCs was performed in a 60 mm petri dish as shown in Figure 3.5. 3yr and WI38 human fibroblasts with a low passage (passage 4) were plated with a

confluence of 1.9×10^6 (~90% of confluence) and cultured with fibroblasts DMEM medium. Infection with viral medium containing OSKM particles was performed using the cellular reprogramming protocol detailed in materials and methods, briefly using a confluence of 85/90% of 293T human cell in a 100 mm petri dish (confluence used was 4×10^6) and 50µl of X-tremeGENE™ 9 transfection reagent diluted in 600µl of DMEM for viral medium production. 5ml of viral medium was added to each 60 mm petri dish and incubated at 37°C overnight. A second infection was performed in the next day (24h infection, Figure 3.5). Both human fibroblasts lines were expressing GFP. The average of GFP expression in WI38 cells was 68.2% and the average of 3yr was 72.4% (Annex 1, Figure 6.3).

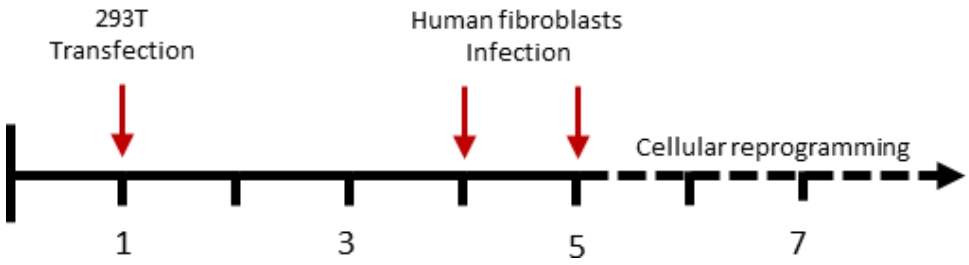
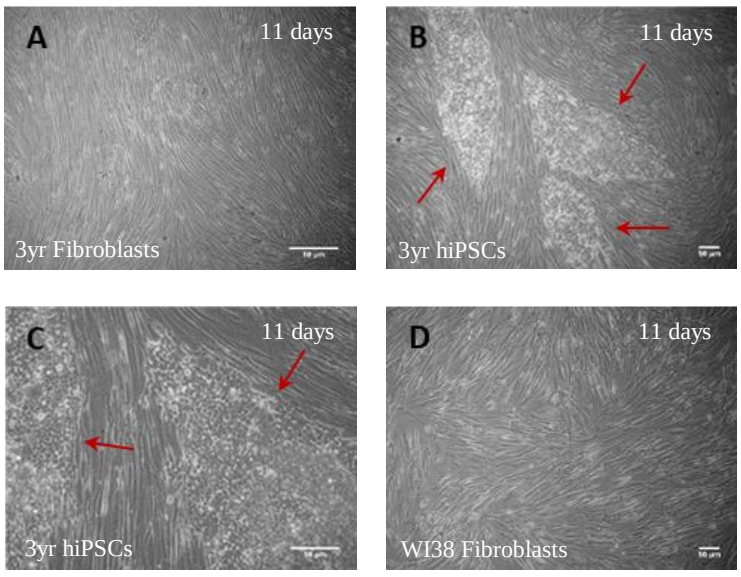


Figure 3.5 - Illustration of cellular reprogramming protocol performed in 3yr and WI38 human fibroblasts with 293T transfection performed in day 1 and human fibroblasts infection in day 4 and 5.

The first hiPSCs clones started to appear 11 days after first infection of cells (Figure 3.6), being a good indication that the cellular reprogramming was optimized for the specific cell lines and experiments.

It was also tested in both human cell lines a lower concentration of viral medium (confluence of 293T was reduced to a quarter) to observe the impact in cellular reprogramming efficiency. In fact, 28 days after first infection, there were no signs of hiPSCs. Although cells infected with high titer viral medium generated hiPSCs clones. This indicates that viral medium concentration is an important criteria to take into account for an efficient cellular reprogramming.



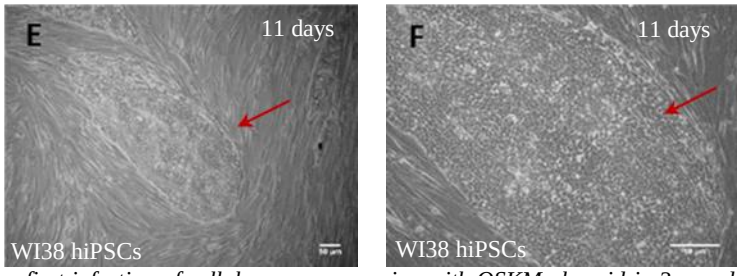


Figure 3.6 - 11 days after first infection of cellular reprogramming with OSKM plasmid in 3yr and WI38 human fibroblasts (passage 4). (A) 3yr human fibroblasts control, amplification x4. (B) 3yr hiPSCs clones reprogrammed amplification x4. (C) 3yr hiPSCs clones reprogrammed, amplification x10. (D) WI38 human fibroblasts control, amplification x4. (E) WI38 hiPSCs clones reprogrammed, amplification x4. (F) WI83 hiPSCs clones reprogrammed, amplification x10. Scale bar was set for 50µm. Images acquired through the Carl Zeiss PrimoVert microscope. Images were treated posteriorly by using Fiji software.

Due to the fact that this assay was only an optimization of cellular reprogramming, no quantification of the number of hiPSCs generated was performed. However, hiPSCs were characterized to confirm its characteristics. First, was performed a passage of hiPSCs using the EDTA protocol in order to isolate the hiPSCs from the fibroblasts. To pass hiPSCs we used two protocols. The protocol using EDTA was used when needed to maintain a hiPSCs clone structure, being more stable and inducing almost no differentiation. The protocol using trypsin was used when needed to detach hiPSCs into single clones.

The hiPSCs isolated were expanded in 6 and 12 well plated pre-coated with Matrigel™ and maintained in mTeSR™1 medium, changed every day. Figure 3.7 show the reprogrammed hiPSCs.

Until characterization, hiPSCs were maintained in culture using the protocols described in the human iPSCs section of materials and methods. It's crucial for hiPSCs to have the right conditions of confluence and medium to maintain their pluripotent properties, otherwise hiPSCs start to differentiate into other germ layers cells. However, as hiPSCs don't survive in single cell, the use of the Rho-associated kinase inhibitor, also known as ROCK inhibitor, was extremely necessary to maintain hiPSCs alive by increasing dissociated human stem cells survival through apoptosis inhibition [36], being added to the first medium added after passing hiPSCs.

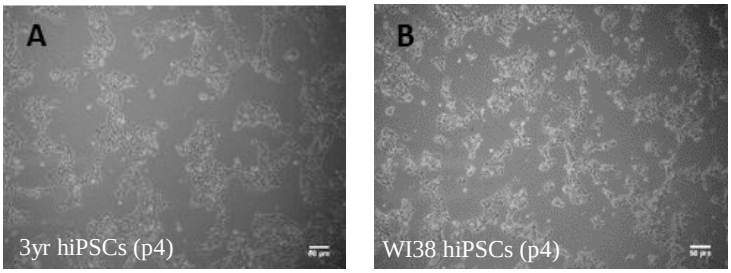


Figure 3.7 - Reprogrammed hiPSCs isolated from feeders after using passing EDTA protocol. (A) Reprogrammed hiPSCs from 3yr human fibroblasts (passage 4º) (B) Reprogrammed hiPSCs from WI38 human fibroblasts (passage 4). Scale bar was set for 50µm. Images acquired through the Carl Zeiss PrimoVert microscope, amplification x4. Images were treated posteriorly by using Fiji software.

The first step of hiPSCs characterization was performing an alkaline phosphatase staining done by using the alkaline phosphatase detection protocol described in the materials and methods section.

As observed in Figure 3.8, in both WI38 and 3yr hiPSCs (A, B, D and E), after alkaline phosphatase treatment, was possible to observe a purple color on the location of hiPSCs clones indicating an alkaline phosphatase activity. 3yr hiPSCs (D and E) were with a high confluence, showing a high alkaline phosphatase activity comparatively to WI38 hiPSCs. Both C and F wells were

not treated with AP staining kit, not showing any purple color as expected. These results partially confirm hiPSCs identity.

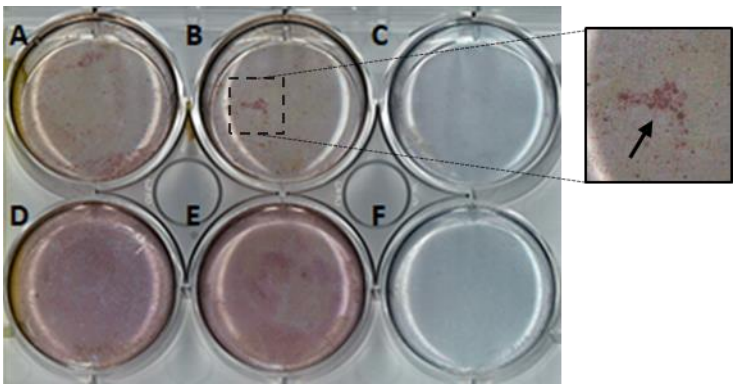


Figure 3.8 - hiPSCs after alkaline phosphatase treatment. A and B - WI38 hiPSCs with low confluence after AP treatment; C - WI38 hiPSCs with low confluence without AP treatment; D and E - 3yr hiPSCs with high confluence after AP treatment; F - 3yr hiPSCs with high confluence without AP treatment.

However, an alkaline phosphatase treatment was not enough to confirm the stem cell properties of hiPSCs. Therefore Immunofluorescence was performed to support the positive results from AP staining by detecting the presence of specific pluripotency markers using specific antibodies. Cells were stained with DAPI (Sigma-Aldrich) and human antibodies against Nanog (Thermo Fischer Scientific) and Sox2 (R&D System) proteins. However, no expression of human Nanog and human Sox2 were observed in both human cell lines. Normally, this indicates a non-pluripotent characteristics in cells, still an alkaline phosphatase activity was previously confirmed.

To confirm exactly whether cells were pluripotent a teratoma formation assay was performed to test the capacity of the reprogrammed hiPSCs to differentiate into the three germ layers in vivo. hiPSCs were injected subcutaneously into dorsal flanks of NSG mice (iMM rodent facility), mice that has no immunity system making them a good strain for human stem cells engraftments. The teratoma formation was observed and dissected 3 weeks post-injection. The teratomas that formed were fixed in formalin and are being currently analyzed in the histology and pathology unit at iMM, for the presence of the 3 germ layers. Although histology results weren't released yet. The teratomas obtained were identical as demonstrated in Figure 3.9.

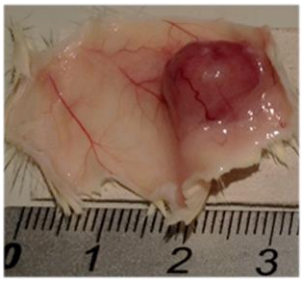


Figure 3.9 - Representative teratoma formed after injection of hiPSCs

3.2. Aging as a barrier for cellular reprogramming

With the cellular reprogramming protocol fully optimized we advance to test the impact of aging in the reprogramming efficiency of human cells.

A RT-qPCR was performed using the cells infected with viral medium containing OSKM particles to determine the levels of expression of the human Oct4 (hOct4) of WI38 and 3yr human fibroblasts with low passage (passage 4) and high passage (passage 7). Samples were prepared by following the RT-qPCR protocol detailed in materials and methods. The RT-qPCR results obtained for hOct4 levels are presented in Figure 3.10.

These results show an increasing of hOct4 levels in both human fibroblasts comparatively with their controls and a high expression of WI38 comparatively with 3yr human fibroblasts, as previously observed. Interestingly, the human cells with a high passage (p7) presented a much higher expression of hOct4, comparatively with human cells with a low passage (p4). This results indicates that transduction using OSKM plasmid was working efficiently.

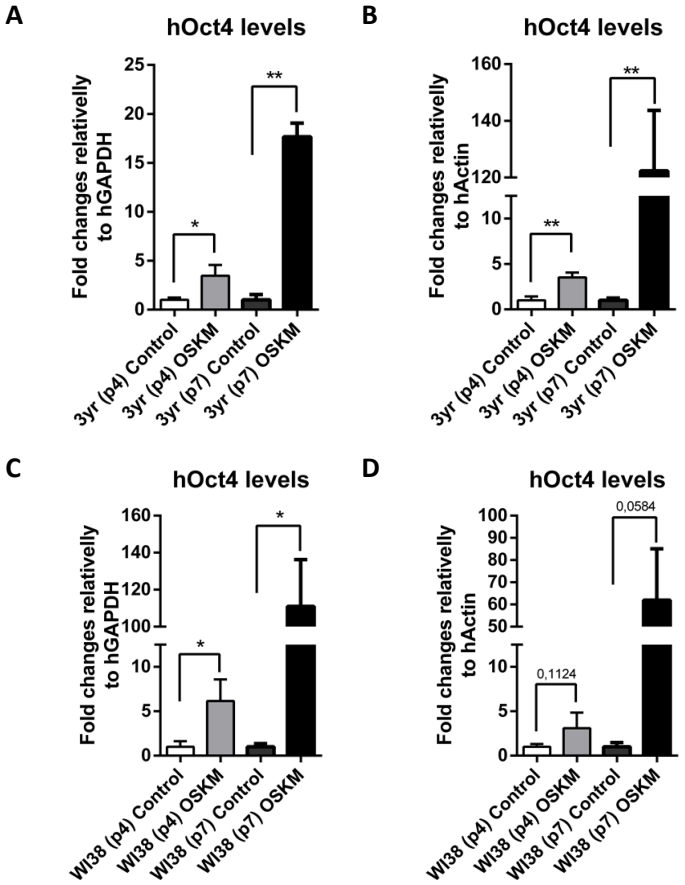


Figure 3.10 - RT-qPCR results of hOCt4 expression levels of WI38 and 3yr human fibroblasts with passage 4 (p4) and passage 7 (p7) presented as $\Delta\Delta C_t$ normalized using non OSKM lentiviral transduced cells as control and hGAPDH and hActin as housekeeping gene. (A and B) WI38 human fibroblasts (p4 and p7) hOct4 levels using housekeeping hGAPDH (A) and hActin (B). (C and D) 3yr human fibroblasts (p4 and p7) hOct4 levels using housekeeping hGAPDH (C) and hActin (D). P-value ≤ 0.05; p-value ≤ 0.01**; p-value ≤ 0.001***.*

To test the impact of aging in the reprogramming process, WI38 and 3yr human fibroblasts with a high passage (passage 7), were plated in a 60 mm petri dish with a confluence of *75%* and cultured with supplemented DMEM medium (see materials and methods). Confluence used was lower due to the difficult of cells with high passage to grow in culture. Cells were infected with 5ml of viral

medium containing OSKM particles produced using an 85/90% of 293T human cell confluence in a 100 mm petri dish (cells plated: 4×10^6) and 50µl of X-tremeGENE™ 9 transfection reagent diluted in 600µl of DMEM. A second infection at 24 hours was performed. The expression of GFP was confirmed having approximately the same efficiency rate between both cell lines with a high passage. (WI38 (p7) = 65.6%; (p7) = 67.8%), confirming a successful infection. The results are presented in annex 1, Figure 6.4. 31 days post-infection, no hiPSCs were detected (Figure 3.11) on the human cells with high passage (passage 7).

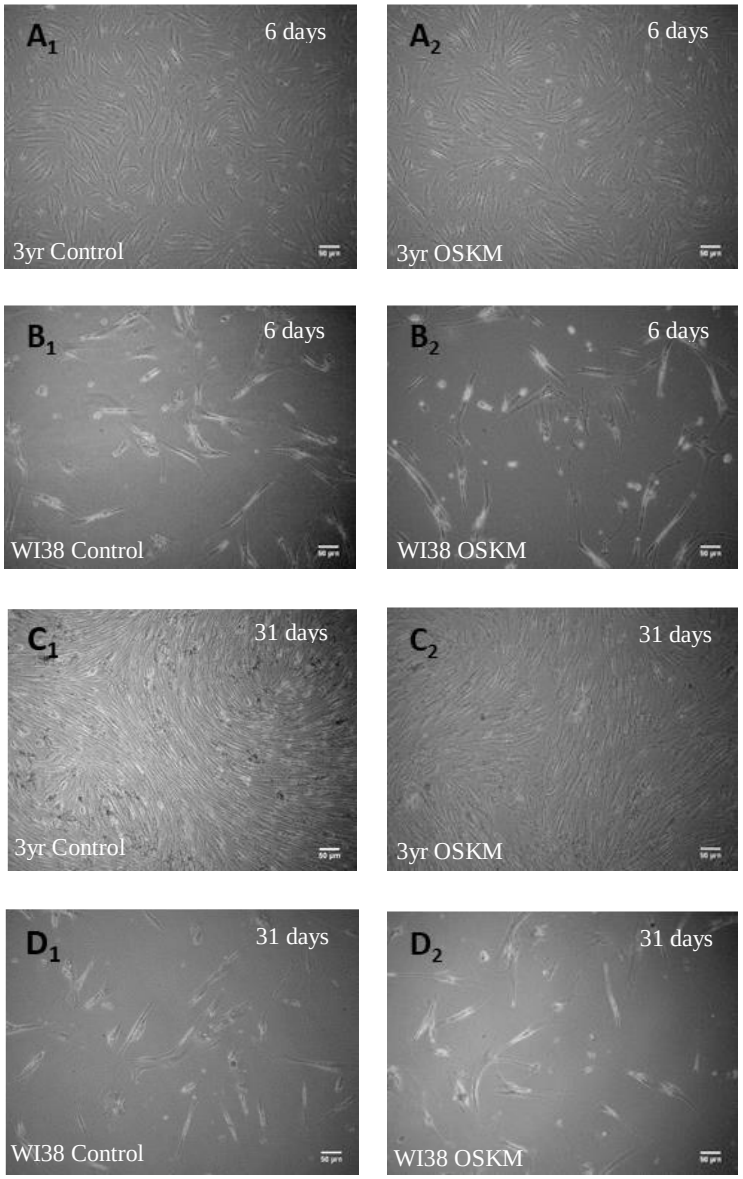


Figure 3.11 –Cellular reprogramming protocol performed in WI38 and 3yr human fibroblasts with high passage (p7). (A and B) 6 days after first infection of 3yr and WI38 human cells lines with a high passage (passage 7) using the optimized cellular reprogramming protocol. (A_1) 3yr human fibroblasts negative control. (A_2) 3yr human fibroblasts transduced with OSKM. (B_1) WI38 human fibroblasts negative control. (B_2) WI38 human fibroblasts transduced with OSKM. (C and D) 31 days after first infection of 3yr and WI38 human cells lines with a high passage (passage 7) using the optimized cellular reprogramming protocol. (C_1) 3yr human fibroblasts negative control. (C_2) 3yr human fibroblasts transduced with OSKM. (D_1) WI38 human fibroblasts negative control. (D_2) WI38 human fibroblasts transduced with OSKM. Scale bar was set for 50µm. Images acquired through the Carl Zeiss PrimoVert microscope, amplification x4. Images were treated posteriorly by using Fiji software.

Despite both human fibroblasts cell lines with high passage (p7) had a much higher expression of hOct4, comparatively with the human cell lines with low passage (p4), no hiPSCs generated after performing the cellular reprogramming protocol, indicating that the number of passages in culture could have an important impact in the cellular reprogramming efficiency. This happened due to the fact that every cell passaging increase the risk of occurring DNA damage, mutations and also the change of some cell line's characteristics giving rise to an alteration in morphology, response to stimuli, growth rates, protein expression and also transfection efficiency of the cells[43]. In fact, was also observed an alteration in morphology of human cells with passage 7 comparatively with human cells with a lower passage (passage 4), suggesting a difference in cell line properties between different passages.

According to Leonard Hayflick and Paul Moorhead human cells have a limit number of divisions in culture, which can vary between cell types and species. Each cell division can induce telomere shortening that results in replicative senescence. Senescence can also be induced by random molecular damage, oxidative stress and DNA damage, during aging. Senescence can be accumulated in some tissue contributing to organ dysfunction[44,45].

This suggests that due to their high number of passages in culture WI38 and 3yr experience telomere shortening, leading to senescent phenotype. Senescence is related with the reprogramming process, indeed it was previously shown that the presence of senescence in human mesenchymal cells decreased their differentiation potential by the accumulation of oxidative stress and the dysregulation of key differentiation regulatory factors [46].

We also observed a reducing of confluence over time which happened due to an increase of cell death after cellular reprogramming protocol was performed. The increase of cell death could be due to the stress induced by the reprogramming protocol in aged/senescent cells that are more susceptible to DNA damage and mutations which could lead to cell death.

To analyze how aging would influence the cellular reprogramming efficiency in human cells, all hiPSCs generated from 3yr and WI38 with low passage (passage 4) were counted between a period of 25 and 30 days after the first infection.

Table 3.1 represents an overview of the number of hiPSCs reprogrammed in either human fibroblasts cell lines.

Table 3.1 - Number of hiPSCs clones reprogrammed from WI38 and 3yr human fibroblasts in 60 mm petri dish with a low passage (passage 4) following the non-optimized and the optimized cellular reprogrammed protocol using OSKM plasmid.

	Before cellular reprogramming protocol optimization		**After cellular reprogramming protocol optimization**
	Assay 1	**Assay 2**	**Assay 3**
3yr hiPSCs	1	0	39
WI38 hiPSCs	7	5	39

From the hiPSCs counting was possible to observe a notorious difference in the amount of hiPSCs generated from 3yr and WI38 human fibroblasts before the protocol was optimized. WI38 generated 12 hiPSCs clones and 3yr generated only 1 clone, suggesting that aging was acting as a barrier for cellular reprogramming of human cells; at least using a no optimized cellular reprogramming protocol.

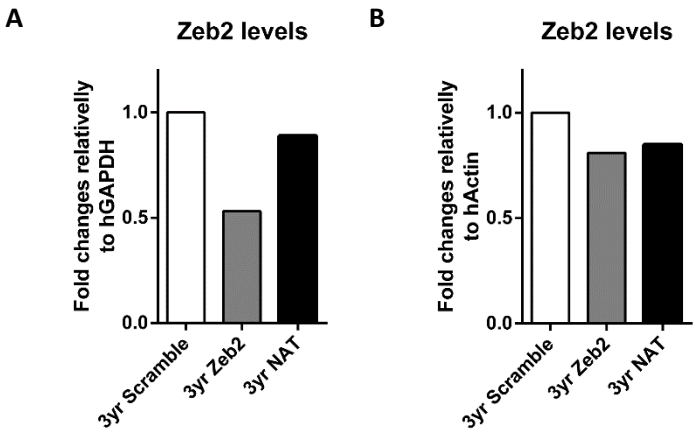
However, after the optimized cellular reprogramming protocol was achieved, no difference between 3yr and WI38 was observed. Both lines generated 39 hiPSCs clones, suggesting that after achieving a cellular reprogramming protocol fully optimized for the specific cell lines and experiment, aging don't have a large influence in cellular reprogramming efficiency in human cells.

3.3. Role of the lncRNA Zeb2NAT on cellular reprogramming of human cells

The initial hypothesis idea was to test if the results from mice experiments were paralleled in human cells, but as the results from the assay intended to test aging as barrier showed a non-influence of aging in cellular reprogramming efficiency, not been possible to confirm that age acts as a barrier in cellular reprogramming of human cells, we focused on understanding exactly the role of the lncRNA Zeb2NAT on cellular reprogramming of human cells. In order to verify the Zeb2NAT role on cellular reprogramming of human cells, a new cellular reprogramming assay followed by LNAs transfection was performed to downregulate the expression of Zeb2 or Zeb2NAT. The hypothesis suggested, with based on results from mice experiments, was that the downregulation of Zeb2 and Zeb2NAT (NAT) would increase the efficiency of cellular reprogramming, improving the quickness and the quality of the hiPSCs generated.

However, before this hypothesis was tested a RT-qPCR was performed following the RT-qPCR protocol to measure the expression levels of lncRNA Zeb2 and Zeb2NAT in both human fibroblasts lines. Downregulation of Zeb2 and Zeb2NAT was performed using LNA GapmeRs.

The expression levels of both lncRNAs Zeb2 and Zeb2NAT are shown in Figure 3.12 and annex 1, Figure 6.5.



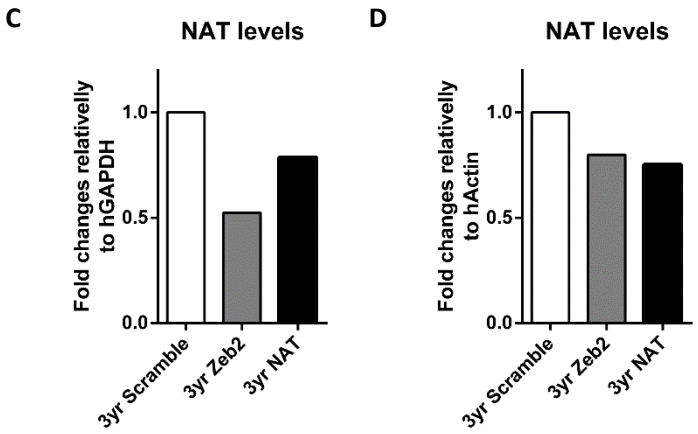


Figure 3.12 - Results of downregulation efficiency of hZeb2 and hNAT expression level in 3yr human fibroblasts presented as $\Delta\Delta C_t$ normalized using non downregulation of lncRNAs Zeb2 and Zeb2NAT as control and HGAPDH and hActin as housekeeping gene. (A and B) ZEB2 expression levels in 3yr cells with knockdown of Zeb2 and (A) and hActin (B) as housekeeping (C and D) 3yr Zeb2, zeb2NAT and scramble levels compared to control values using hGAPDH (C) and hActin (D) as housekeeping.

The RT-qPCR results show that the cells with the downregulation of Zeb2 have a decrease of both Zeb2 and NAT genes comparatively with the cells with the downregulation of NAT, using a non-specific control (scramble) as a reference. This shows that LNA transfection is working properly. However, was expected that cells with NAT downregulation would express a lower levels of NAT comparatively with cells with Zeb2 downregulation.

To verify the enhancing on cellular reprogramming efficiency by downregulating lncRNAs Zeb2 and Zeb2NAT a cellular reprogramming assay followed by LNAs GapmeRs transfection was performed in a 12 well plate (Figure 3.13).

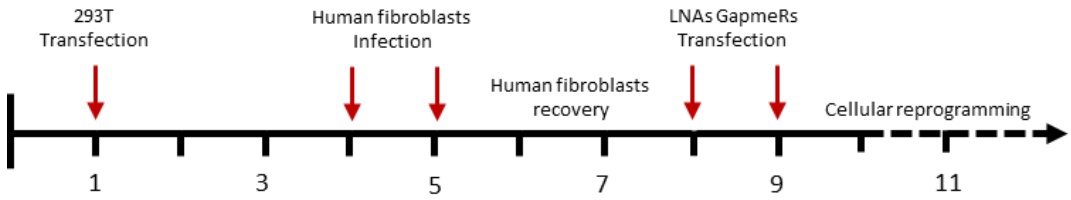


Figure 3.13 – Cellular reprogramming protocol with LNAs GapmeRs transfection protocol in 3yr and WI38 human fibroblasts. Cellular reprogramming protocol started in day 1 with 293T transfection and ended in day 5 with second infection of human fibroblasts. LNAs GapmeRs transfection protocol was performed in day 8 and 9.

3yr and WI38 human fibroblasts with a low passage (passage 4) were plated with a confluence of 3×10^5 (~90% of confluence) and cultured with supplemented DMEM medium. Infection was performed by using the optimized cellular reprogramming protocol. 1ml of viral medium was added to each well for the first and the second infection with an incubation period of 24 hours each at 37°C. Both human fibroblasts lines expressed GFP (WI38 (p4) = 61.0%; 3yr (p4) = 67.8%). The results are presented in annex 1, Figure 6.6. Cells were maintained in DMEM for a period of two days after second infection to allow them to recover from the stress-induced infection.

Afterward, a knockdown of Zeb2 and Zeb2NAT was performed in cells already infected by using LNA GapmeRs as described in LNA GapmeRs in materials and methods section.

14 days after first infection, only control and scramble cells were reprogramming. 19 days after, cells with Zeb2 downregulation started to reprogram (Figure 3.14). However, no hiPSCs were observed in both cell lines with the Zeb2NAT knockdown.

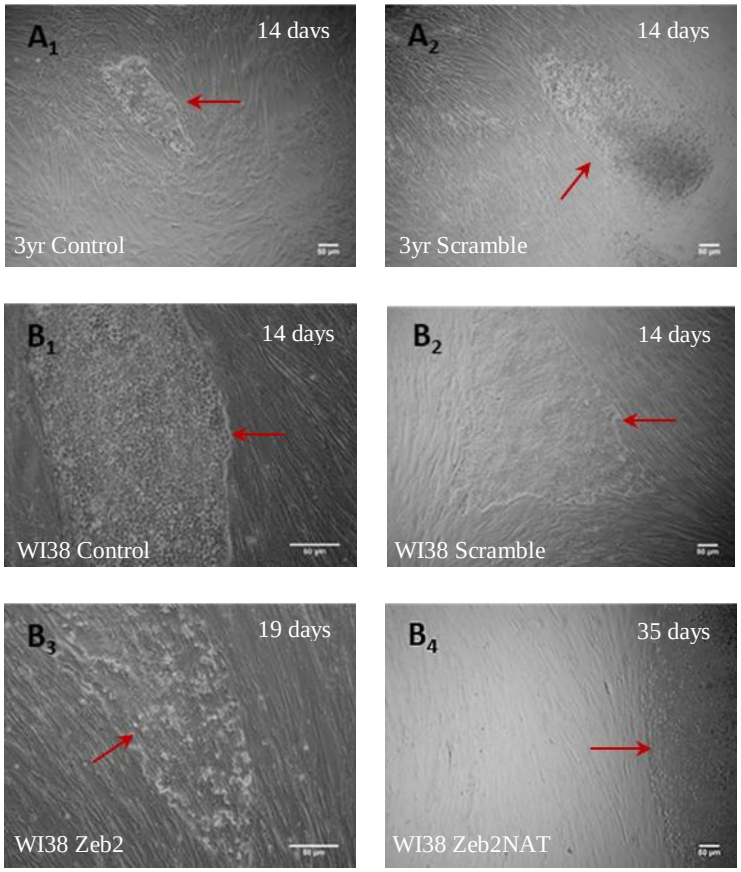


Figure 3.14 – HiPSCs generated after performing the cellular reprogramming protocol and the LNA transfection protocol in both human cell lines. 3yr (A) and WI38 (B) human fibroblasts with low passage (p4) after knockdown of Zeb2 and Zeb2NAT. (A_1) 3yr human fibroblasts negative control for LNA transfection. Started to reprogram 14 days after first reprogramming infection, amplification x4. (A_2) 3yr human fibroblasts transfected with scramble. Started to reprogram 14 days after first reprogramming infection, amplification x4. (B_1) WI38yr human fibroblasts negative control for LNA transfection. Started to reprogram 14 days after first reprogramming infection, amplification x10 (B_2) WI38yr human fibroblasts transfected with scramble. Started to reprogram 14 days after first reprogramming infection, amplification x4. (B_3) WI38 human fibroblasts after knockdown of Zeb2. Started to reprogram 19 after first reprogramming infection, amplification x10. (B_4) WI38 human fibroblasts after knockdown of Zeb2NAT. Started to reprogram 35 days after first reprogramming infection. Scale bar was set for 50µm. Images acquired through the Carl Zeiss PrimoVert microscope. Images were treated posteriorly by using Fiji software.

14 days after the first transfection in both cell lines, hiPSCs started to appear in both control and Scramble conditions, as expected. However, control condition generated more hiPSCs in both cell lines. No hiPSCs were generated in the 3yr and WI38 cells with the downregulation of Zeb2 and zeb2NAT after 14 days. Only 19 days after the first transfection, W38 human fibroblasts with the Zeb2 downregulation started to reprogram and the WI38 with the downregulation of the Zeb2NAT started to reprogram only 35 days after the first transfection (Figure 3.14).

This results go against what was expected, having in consideration that mouse cells with knockdown of Zeb2 and Zeb2NAT reduced the time needed to form miPSCs. It seems that downregulation of both Zeb2 and Zeb2NAT are, somehow, blocking the capacity to reprogram, human cells. Could be possible that this blocking was being caused by the transfection itself and not specifically caused by the knockdown of Zeb2 and Zeb2NAT. In fact, was already demonstrated that the use of Lipofectamine (in this case, RNAiMAX reagent) activate stress genes affecting cell cycle regulation and/or metabolic signaling [47]. However, if the transfection was only the reason to a lower reprogramming efficiency, it would be expected to observe a decrease and a delay of the clones generated in cells transfected with Scramble, which was not the case. In fact, cells with Scramble transfection took the same time as the control to generate hiPSCs, and the difference in the number of

clones generated was not significant to consider the LNAs transfection as the only responsible for a lower cellular reprogramming efficiency. If not, cells with Zeb2 an Zeb2NAT downregulation would have the same response as cell with Scramble, suggesting that, downregulation of Zeb2 and Zeb2NAT, somehow, are blocking and delaying the generation of hiPSCs through cellular reprogramming.

To calculate the cellular reprogramming efficiency was performed a quantification of the number of hiPSCs generated 28 days after the first reprogramming infection, shown in Table 3.2.

Table 3.2 - Quantification of number of hiPSCs clones 28 days after first cellular reprogramming infection and 25 days after 1º LNA transfection.

	ZEB2	NAT	Scramble	Negative control	Concentration	GFP expression
3yr (p4)	0	0	10	12	3×10^5	67,8%
WI38 (p4)	4	0	8	10	3×10^5	61%

The Table 3.3 shows the rate of cellular reprogramming efficiency for both human cell lines (WI38 and 3yr) after downregulation of Zeb2 and Zeb2NAT and using Scramble and negative control as the reference. The results shows that the negative control of both WI38 and 3yr have the best reprogramming efficiency (with 0.006%/0.005% of reprogramming rate for 3yr/WI38 for negative control and 0.005%/0.004% for 3yr/WI38 for scramble condition). The cells after Scramble transfection, shows a slight decrease in the efficiency rate, indicating that LNA transfection could be affecting the cellular reprogramming efficiency as explained above. However, none of the WI38 or 3yr with Zeb2 and Zeb2NAT knockdown reprogrammed with the expectation of WI38 with Zeb2 knockdown that reprogramed but still with a low efficiency (0.002%) comparatively with negative control and scramble. This supports the idea that the Zeb2 and Zeb2NAT downregulation can act as barrier for cellular reprogramming. Despite the influence that the Transfection of LNAs has in the efficiency of cellular reprogramming protocol in the human model.

Table 3.3 –Reprogramming efficiency (%) of 3yr and Wi38 human cells with passage 4, 28 days after first cellular reprogramming infection and 25 days after 1º LNA transfection.

	ZEB2 Reprogramming rate	NAT Reprogramming rate	Scramble Reprogramming rate	Negative control Reprogramming rate
3yr (p4)	0%	0%	0.005%	0.006%
WI38 (p4)	0.002%	0%	0.004%	0.005%

Before hiPSCs characterization, hiPSCs were picked directly with a filter tip and isolated from the fibroblasts and the others hiPSCs. The isolated hiPSCS clones were expanded separately in 6 and 12 well plates pre-coated with Matrigel™ and maintained in mTeSR™1 medium, changed every day.

Since, the AP standing results don't allow to firmly state the pluripotent proprieties of the hiPSCs, an immunofluorescence was the first hiPSCs characterization to be performed. Cells were stained with DAPI and human antibodies against Nanog and Sox2.

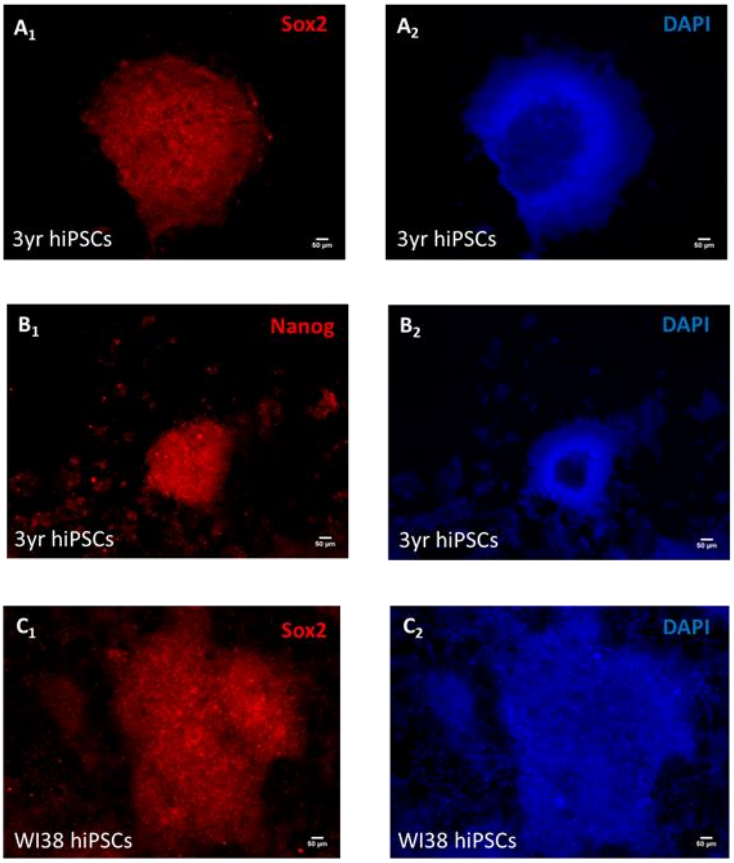


Figure 3.15 - Immunofluorescence performed in 3yr control (A and B) and WI38 control (C) hiPSCs. Cells stained with DAPI in blue and human antibodies against Nanog and Sox2 in red. (A_1) 3yr hiPSCs stained with human antibody agains Sox2 in red. (A_2) 3yr hiPSCs stained with DAPI in blue. (B_1) 3yr hiPSCs stained with human antibody agains Nanog in red. (B_2) 3yr hiPSCs stained with DAPI in blue. (C_1) WI38 hiPSCs stained with human antibody agains Sox2 in red. (C_2) Wi38 hiPSCs stained with DAPI in blue. Scale bar was set for 50µm. Images acquired through the Carl Zeiss Axiovert 200M fluorescence microscope. Images were treated posteriorly by using Fiji software.

The immunofluorescence results presented in Figure 3.15 showed that hiPSCs generated from WI38 and 3yr control human fibroblasts expressed the pluripotency markers, comparing to negative control However, WI38 hiPSCs didn't expressed Nanog. Besides that, no substantial difference was observed between both hiPSCs lines.

The reprogrammed hiPSCs from this experiment are being expanded in a 100mm culture petri dish in order to perform a teratoma formation assay to confirm the capacity of reprogrammed hiPSCs to differentiate into the three different germ layers in vivo, as is expected.

Due to the fact that knockdown of Zeb2 and Zeb2NAT was not enhancing reprogramming efficiently (was even blocking the capacity to reprogram), with the generation of a really low quantity of hiPSCs clones, going against what was already observed in assays performed with mice cells, two hypothesis were suggested (A) The order of reprogramming protocol could somehow affect the efficiency of cellular reprogramming inducing some stress in cells, already proven to happen in cellular reprogramming of mice cells (data not shown); (B) The lncRNA Zeb2 and Zeb2NAT could have a different mechanisms in human cells.

The hypothesis (A) was tested by changing only the order of the protocol. The knockdown of both Zeb2 and Zeb2NAT occurred first. Afterwards WI38 and 3yr cells were infected with viral medium specific to induce the four transcriptional factors (Figure 3.16). However, due to the fact that this assay was performed recently, no results are still available.

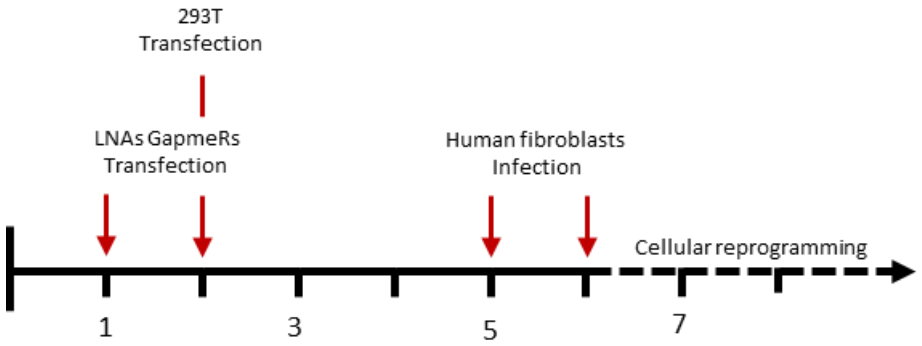


Figure 3.16 - Illustration of LNAs GapmeRs transfection protocol with cellular reprogramming protocol performed afterwards in 3yr and WI38 human fibroblasts. LNAs GapmeRs transfection protocol was performed in day 1 and 2. Cellular reprogramming protocol started in day 2 with 293T transfection and ended in day 6 with second infection of human fibroblasts.

During the optimization of cellular reprogramming protocol needed to perform the experiments using both human cell lines, WI38 and 3yr human fibroblasts, was notorious the influence that certain factors, such as a high confluence of both human fibroblasts and 293T and a good concentration of the viral medium had to increase the efficiency of cellular reprogramming.

The first objective was to analyze how aging could affect reprogramming efficiency. Was already proven that aged mice cells had a lower preprogramming efficiency, comparatively with younger mice cells, however, no tests were performed using human cells to validate this hypothesis. Initially, during the optimization of reprogramming protocol was possible to observe that WI38 cells generated more hiPSCs comparatively with 3yr cells, both cells lines with a low passage. In the total, WI38 generated 12 hiPSCs while 3yr generated only 1.

These preliminary results were indicating a probable reduction of cellular reprogramming induced by aging. However, when reprogramming the same human cell lines with a low passage using a fully optimized cellular reprogramming protocol, no difference in the number of hiPSCs generated in both human cells was observed. Both cell lines reprogramed exactly the same 39 hiPSCS clones after a period of 25-30 days. This results were against what was expected, since aging was observed to act as a cellular reprogramming barrier in mice cells, however, this results suggested that aging, somehow, can act as a barrier while the cellular reprogramming protocol is not fully adapted. When a cellular reprogramming assay was performed with the fully optimized protocol there was no evidence that aging was reducing the cellular reprogramming efficiency, otherwise would be expectable to observe a lower number of hiPSCs generated from 3yr human fibroblasts comparatively with WI38 human fibroblasts.

In addition to aging, it was also tested how the number of passages of cell in culture could influence the efficiency of hiPSCs generation. Interestingly, human cells with a high passage (p7) expressed high levels of hOct4, comparatively with human fibroblasts with low passage (p4), however when WI38 and 3yr human fibroblasts with a high passage (passage 7) were used no hiPSCs were detected, even after 60 days, contrary to what was observed when reprogramming WI38 and 3yr with a low passage (passage 4), that started to reprogram into hiPSCs 11 days after first infection. hiPSCs

characterization show that, there were not evident signs that WI38 hiPSCs had a better quality, comparatively with 3yr hiPSCs, as expectable due to their difference in age.

The initial idea was to test if aging was acting as barrier for cellular reprogramming based in the experiments with mice cells performed in our lab, however, no evidence were found supporting this idea. In fact, the results suggested that the number of passages had a higher impact in the cellular reprogramming efficiency that aging. With no results showing the effects of aging in the cellular reprogramming, we focused on understanding the lncRNA Zeb2NAT role on cellular reprogramming of human cells. The new hypothesis suggested, with based on results from mice experiments, was that the knockdown of Zeb2 and Zeb2NAT would increase the efficiency of cellular reprogramming, improving the quickness and the quality of the hiPSCs generated.

After cellular reprogramming protocol was accomplish in WI38 and 3yr cells with a low passage, a downregulation of Zeb2 and Zeb2NAT was performed. 14 days after first infection, both control and Scramble cells started to reprogram with a higher number of clones generated in control, suggesting that LNA transfection was inducing some kind of stress in cells reducing the efficiency of cellular reprogramming. We observed that with scramble, a reduction of 20% in generated hiPSCs was achieved.

Interestingly, and contrary to mice cells, we observed that cells with the downregulation of Zeb2 and Zeb2NAT had a lower hiPSCs generation. In fact, only WI38 cell with the knockdown of Zeb2 and Ze2NAT reprogrammed, however, with a temporary delay since only after 19 and 35 days, after first infection, WI38 human fibroblasts started to reprogram, correspondingly. No hiPSCs were generated in 3yr cells with the downregulation of Zeb2 and Zeb2NAT.

This suggest that not also the LNA transfection reduce the efficiency of cellular reprogramming but the Zeb2 and Zeb2NAT downregulation have a great impact in the efficiency rate, blocking and delaying, of the reprogramming of human cells. This can mean that the lncRNA Zeb2 can have a different mechanisms in human cells, from what was observed in mice cells or the need to optimize the cellular reprogramming protocol and the LNA transfection protocol, in order to adapt both in the same experiment.

3.4. NORAD affects chromosomal stability of iPSCs after DNA damage

Was also intended to understand how the lncRNA NORAD could affect stem cells stability, viability and pluripotent characteristics. As demonstrated by Mendell and colleagues, the absence of NORAD triggers chromosomal instability in the human colon cell line through a release of PUMILIO proteins. Another objective was also to try to obtain stem cells more stable and with better quality by regulating PUM2 protein through the downregulation and upregulation of lncRNA NORAD. As already know, that PUM2 is required for the germ line stem cells maintenance, two hypothesis were suggested: (A) if the NORAD was activated after DNA damage this would induce a premature differentiation in germ line stem cells since the NORAD would sequester the PUM2 from the cells. (B) if the NORAD deletion was performed, PUM2 would not be sequestered by NORAD after DNA damage, allowing PUM2 proteins to maintain germ cells lines, preventing a premature differentiation after DNA damage.

To test NORAD role was used a CRISPR/CAS9 technique to delete NORAD gene in E14 stem cells. For that, two pairs of oligonucleotides, were specifically designed with a mouse homology sequence

from the mice 4.9kb lncRNA NORAD (290097C17rik) (Figure 3.17 - A) and annex 2). The two plasmids used to perform the transfection of E14 for NORAD deletion were prepared using the NORAD plasmid construction protocol described in methods section.

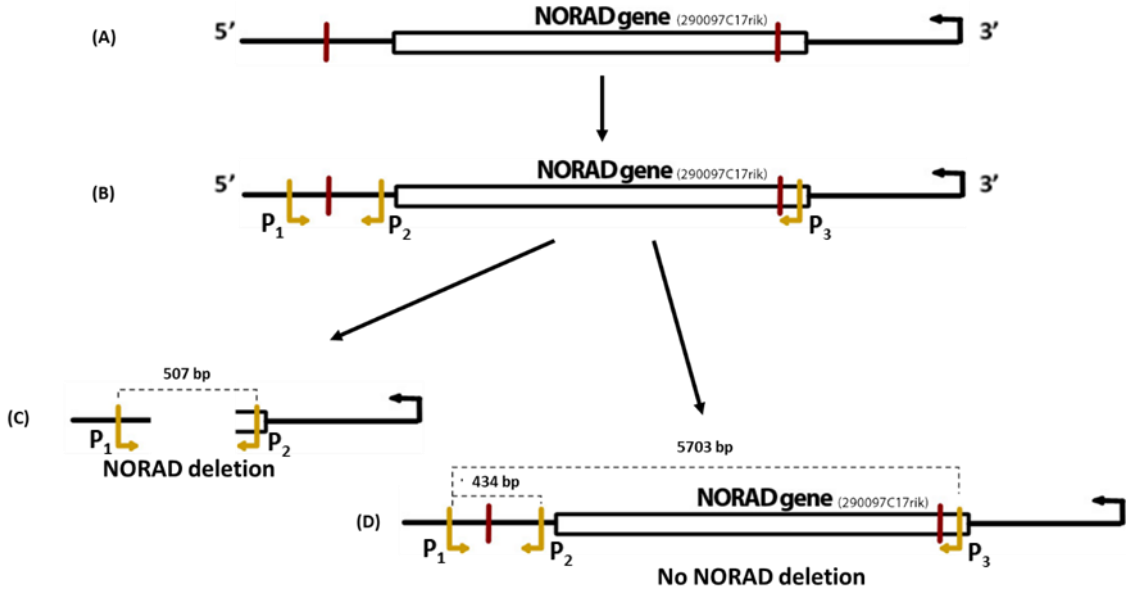


Figure 3.17 - Representation of CRISPR/CAS9 technique used to delete NORAD gene. (A) Norad plasmids in red designed to cut in a specific location for NORAD deletion. (B) Two par of primers in yellow designed to detect if NORAD was successfully deleted through electrophoresis. (C) If NORAD was successfully deleted a band of 507bp in electrophoresis gel would be detected. (D) If NORAD deletion was not accomplished two bands of 434bp and 5703bp would appear in electrophoresis gel.

E14 were plated in a 6 well plate at least 24 hours prior to the transfection. The NORAD transfection was performed as described in materials and methods section, however, two conditions were tested: a transfection with 1µl of each NORAD plasmid and other transfection with 2 µl of each NORAD plasmid. After both transfections were performed, the E14 were maintained in KSR+LIF. Two days after the transfection protocol, 1µg of Puromycin was added to KSR+LIF medium for clone selection.

The clones that survived were picked and expanded. After reached a good confluence, a genomic DNA extraction and a PCR were performed using a lysis buffer and a Taq NEB enzyme, correspondingly, to confirm if the selected clones had the NORAD gene deleted by an electrophoresis on 0,5x TBE with 1% agarose gel stained with Midori Green Advance using two pairs of primers specifically design to detect the cuts from the two pairs of plasmids (Figure 3.17 - B).

However, from all clones that were picked, not a single clone had the NORAD deletion. To be able to delete NORAD gene from E14, another approach was needed.

An electroporation method was performed to delete NORAD gene in E14 stem cells. The electroporation was performed by using electroporation protocol from material and methods, using the following conditions: 3 pulses of 10ms with a voltage of 1400V.

After Puromycin treatment, no E14 clone survived. E14 cells were exposed for 48 hours to the Puromycin treatment, a long period of time, killing all cells with or without the plasmids incorporated.

Another electroporation is needed to be performed, using only at maximum a period of 24 hours of Puromycin treatment, for clone selection. However, due to the fact that this assay is being performed, no results are still available.

4. Conclusion

Although aging affects reprogramming efficiency in mice cells, it seems that does not have a considerable influence in reprogramming human cells into hiPSCs. However, the number of passages, influences the efficiency of cellular reprogramming. The higher the number of passages, the lower will be the cellular reprogramming efficiency.

A good way to increase this efficiency and to improve the quality of hiPSCs, could pass by downregulating the lncRNA Zeb2NAT a natural antisense transcript of the coding gene Zeb2, already proven to be a good way to increase the efficiency of cellular reprogramming in mice cells. However, when using a human model the downregulation of the Zeb2 and the lncRNA Zeb2NAT reduced the cellular reprogramming efficiency, instead of increasing. Although the LNA transfection contributes to reduce the efficiency of the process, it seems that Zeb2 and Zeb2NAT knockdown induce a kind of delay or even a blocking in cellular reprogramming, suggesting a different Zeb2/Zeb2NAT mechanism in the human cells. Another reason for this low efficiency could pass by the need to optimize the LNA transfection protocol, already known to induce some stress in cells. Further experiments are needed to identify the true reason for this loss of efficiency.

It would also be important to analyze the impact that others plasmids, that induce the expression of the four transcriptional factors, and the influence that others cells types and ages could have in cellular reprogramming efficiency. After performing this experiments to tune the cellular reprogramming protocol to its maximum, would be possible to do other experiments more connected with regenerative medicine and therapeutic purposes.

Despite this, we believe this approach constitutes a novel strategy to study the impact of lncRNAs in cellular identity and we anticipate that the output of this proposal will generate important contributions to the aging and cellular reprogramming research field.

Another way to enhance the quality of stem cells could be through the manipulation of lncRNA NORAD. However, due to the fact that NORAD is a novel lncRNA still very little is known, making it an excellent target to be tested. Due to the recent interest of our lab in this lncRNA, only some initial experiments, using CRIPR/CAS9 technique were performed with the objective of testing how stem cells would react in terms of chromosomal stability and what changes would happen to stem cells morphology and its pluripotent properties after NORAD gene deletion. This approach could provide a good way to produce stem cells more stable and with better quality for clinical purposes.

This page intentionally left blank

5. References

1. Natarajan, K. N., Teichmann, S. A. & Kolodziejczyk, A. A. Single cell transcriptomics of pluripotent stem cells : reprogramming and differentiation. *Curr. Opin. Genet. Dev.* **46,** 66–76 (2017).

2. Menon, S., Shailendra, S., Renda, A., Longaker, M. & Quarto, N. An Overview of Direct Somatic Reprogramming : The Ins and Outs of iPSCs. *Int. J. Mol. Sci.* 1–20 (2016).

3. Hemmat, S., Lieberman, D. M. & Most, S. P. An Introduction to Stem Cell Biology. *Facial Plast. Surg.* **26,** (2010).

4. Weger, M., Diotel, N., Dorsemans, A., Dickmeis, T. & Weger, B. D. Stem cells and the circadian clock. *Dev. Biol.* 0–1 (2017).

5. Juty, N. Chickarmane and Peterson (2008), A computational model for understanding stem cells, trophectoderm and endoderm lineage determination. *EMBL-EBI* (2010). Available at: http://www.ebi.ac.uk/biomodels-main/static-pages.do?page=ModelMonth%2F2010-06. (Accessed: 10th December 2016)

6. Chen, A. E. *et al.* Optimal timing of inner cell mass isolation increases the efficiency of human embryonic stem cell derivation and allows generation of sibling cell lines. *Cell Stem Cell* **4,** 103–106 (2012).

7. Byrne, J. A. *et al.* Producing primate embryonic stem cells by somatic cell nuclear transfer. *Nature* **450,** (2007).

8. Takahashi, K. & Yamanaka, S. A decade of transcription factor-mediated reprogramming to pluripotency. *Nat. Publ. Gr.* **17,** 183–193 (2016).

9. Loi, P. *et al.* Sheep : The First Large Animal Model in Nuclear Transfer Research. *liebertpub* **15,** (2013).

10. Tada, M., Takahama, Y., Abe, K., Nakatsuji, N. & Tada, T. Nuclear reprogramming of somatic cells by in vitro hybridization with ES cells. *Elsevier Sci.* 1553–1558 (2001).

11. Cowan, C. A. Nuclear Reprogramming of Somatic Cells After Fusion with Human Embryonic Stem Cells. *Science (80-.).* **1369,** (2012).

12. Blau, H. M. & Chiu, C. Cytoplasmic Activation of Human Nuclear Genes in Stable Heterocaryons. *Cell* **32,** 1171–1180 (1983).

13. Teoh, H. & Cheong, S. Induced pluripotent stem cells in research and therapy. *Malaysian J Pathol* **34,** 1–13 (2012).

14. Takahashi, K. & Yamanaka, S. Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors. *Cell* **2,** 663–676 (2006).

15. Yamanaka, S. Ekiden to iPS Cells. *Nat. Med.* **15,** 1145–1148 (2009).

16. Ebrahimi, B. Tissue and Cell Engineering cell fate : Spotlight on cell-activation and signaling-directed lineage conversion. *Elsevier* **48,** 475–487 (2016).

17. Kulessa, H., Frampton, J. & Al, E. GATA-1 reprograms avian myelomonocytic cell lines into eosinophils, thromboblasts, and erythroblasts. *Genes Dev.* 1250–1262 (1995).

18. Soria-valles, C. & López-otín, C. iPSCs : On the Road to Reprogramming Aging. *Trends Mol. Med.* **22,** 713–724 (2016).

19. Mahmoudi, S. & Brunet, A. Aging and reprogramming : a two-way street. *Curr. Opin. Cell*

Biol. 1–13 (2012).

20. Greive, S. J., Lins, A. F. & Hippel, P. H. Von. Assembly of an RNA-Protein Complex - Binding of NusB and NusE (S10) proteins to boxA RNA nucleates the formation of the antitermination complex involved in controlling rRNA transcription in escherichia coli. *J. Biol. Chem.* **280,** 36397–36408 (2005).

21. Cech, T. R. & Steitz, J. A. Review The Noncoding RNA Revolution — Trashing Old Rules to Forge New Ones. *Cell* **157,** 77–94 (2014).

22. Rna, N., Quinn, J. J. & Chang, H. Y. Unique features of long non-coding RNA biogenesis and function. *Nat. Publ. Gr.* **17,** 47–62 (2016).

23. Flynn, R. A. & Chang, H. Y. Review Long Noncoding RNAs in Cell-Fate Programming and Reprogramming. *Stem Cell* **14,** 752–761 (2014).

24. Wu, C. *et al.* Senescence-associated Long Non-coding RNA (SALNR) Delays Oncogene-induced Senescence through NF90 Regulation. *J. Biol. Chem.* **290,** 30175–30192 (2015).

25. Loewer, S. *et al.* Large intergenic non-coding RNA-RoR modulates reprogramming of human induced pluripotent stem cells. *Nat. Genet.* **42,** (2010).

26. Li, R. *et al.* A Mesenchymal-to-Epithelial Transition Initiates and Is Required for the Nuclear Reprogramming of Mouse Fibroblasts. *Cell Stem Cell* **7,** 51–63 (2010).

27. Lu, X. *et al.* Variations in mesenchymal-epithelial transition-related transcription factors during reprogramming of somatic cells from different germ layers into iPSCs. *J. Genet. Genomics* (2016).

28. Lamouille, S., Xu, J. & Derynck, R. Molecular mechanisms of epithelial–mesenchymal transition. *Nat Rev Mol Cell Biol* **15,** 178–196 (2014).

29. Jesus, B. *et al.* A long non-coding RNA antisense of Zeb2 enhances reprogramming and pluripotency.

30. Pauli, A., Rinn, J. L. & Schier, A. F. Non-coding RNAs as regulators of embryogenesis. *Nat. Publ. Gr.* **12,** 136–149 (2011).

31. Takata, R., Makado, G., Kitamura, A. & Watanabe, H. A novel dual lock method for down-regulation of genes , in which a target mRNA is captured at 2 independent positions by linked locked nucleic acid antisense oligonucleotides. *RNA Biol.* **13,** 279–289 (2016).

32. EXIQON. Locked Nucleic Acid. (2009). Available at: http://www.exiqon.com/lna-technology. (Accessed: 19th June 2017)

33. Ventura, A. NORAD : Defender of the Genome. *Trends Genet.* **32,** 390–392 (2016).

34. Lee, S. *et al.* Noncoding RNA NORAD Regulates Genomic Stability by Sequestering PUMILIO Proteins. *Cell* **164,** 69–80 (2016).

35. Moore, F. L. *et al.* Human Pumilio-2 is expressed in embryonic stem cells and germ cells and interacts with DAZ (Deleted in AZoospermia) and DAZ-Like proteins. *PNAS* **100,** 538–543 (2002).

36. Watanabe, K. *et al.* A ROCK inhibitor permits survival of dissociated human embryonic stem cells. *Nat. Biotechnol.* **25,** 681–686 (2007).

37. Li, H. *et al.* The Ink4/Arf locus is a barrier for iPS cell reprogramming. *Nature* **460,** 2–7 (2009).

38. Yuan, J. S., Reed, A., Chen, F. & Jr, C. N. S. Statistical analysis of real-time PCR data. **12,** 1–12 (2006).

39. Lever, A. M. L. in *Advances in Pharmacology* **48,** 27 (Academic Press, 2000).

40. Connolly, J. B. Lentiviruses in gene therapy clinical research. *Gene Ther.* 1730–1734 (2002).

41. Cribbs, A. P., Kennedy, A., Gregory, B. & Brennan, F. M. Simplified production and concentration of lentiviral vectors to achieve high transduction in primary human T cells. *BMC Biotechnol.* **13,** 1 (2013).

42. Dull, T. O. M. *et al.* A Third-Generation Lentivirus Vector with a Conditional Packaging System. *J. Virol.* **72,** 8463–8471 (1998).

43. ATCC. Passage Number Effects In Cell Lines. (2010). Available at: https://www.atcc.org/~/media/PDFs/Technical Bulletins/tb07.ashx. (Accessed: 10th September 2017)

44. Pedro, J. & Magalhães, D. Mechanisms of Ageing and Development. *Elsevier* 0–1 (2017).

45. Deursen, J. M. Van. The role of senescent cells in ageing. *Nature* **509,** 439–446 (2014).

46. Turinetto, V., Vitale, E. & Giachino, C. Senescence in Human Mesenchymal Stem Cells : Functional Changes and Implications in Stem Cell-Based Therapy. *Int. J. Mol. Sci.* **17,** 1–18 (2016).

47. Fiszer-kierzkowska, A., Vydra, N., Wysocka-wycisk, A. & Kronekova, Z. Liposome-based DNA carriers may induce cellular stress response and change gene expression pattern in transfected cells. *BMC Mol. Biol.* **12,** (2011).

This page intentionally left blank

6. Annexes

Annex 1

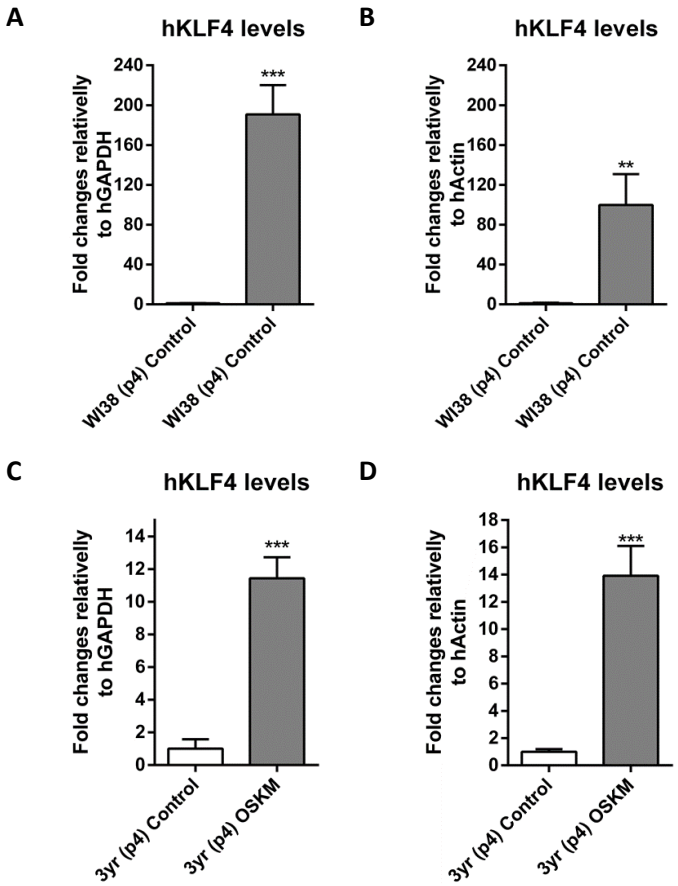
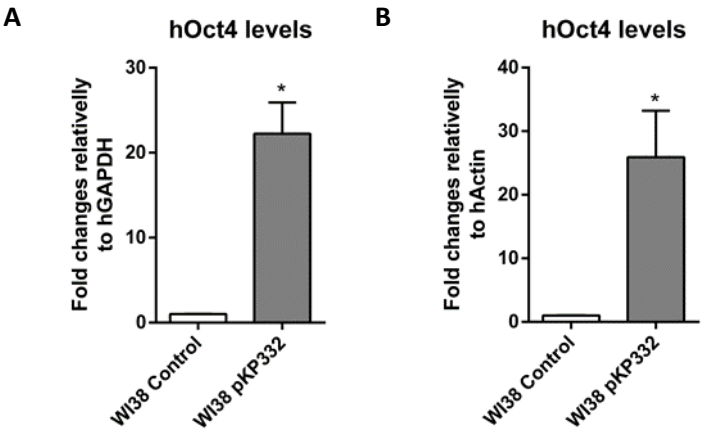


Figure 6.1 - RT-qPCR results of hKLF4 expression levels of WI38 and 3yr human fibroblasts presented as $\Delta\Delta C_t$ normalized using non OSKM lentiviral transduced cells as control and hGAPDH and hActin as housekeeping gene. (A and B) WI38 human fibroblasts hKLF4 levels using housekeeping hGAPDH (A) and hActin (B). (C and D) 3yr human fibroblasts hKLF4 levels using housekeeping hGAPDH (C) and hActin (D). P-value ≤ 0.05; p-value ≤ 0.01**; p-value ≤ 0.001***.*



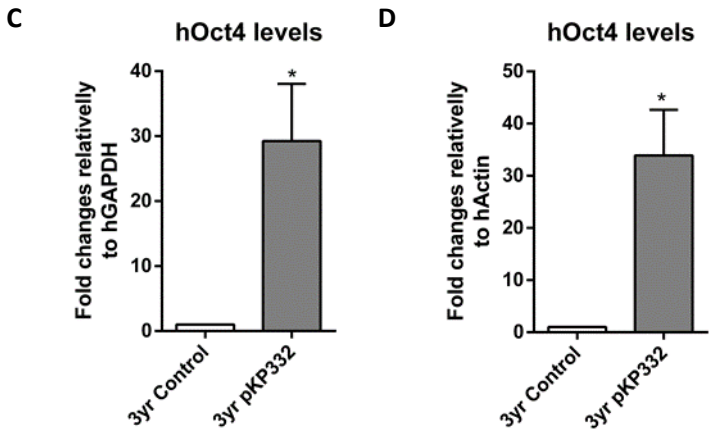


Figure 6.2 - RT-qPCR results of hOct4 expression levels of WI38 and 3yr human fibroblasts (after transduction using pKP332 plasmid) presented as $\Delta\Delta C_t$ normalized using non pKP332 lentiviral transduced cells as control and hGAPDH and hActin as housekeeping gene. (A and B) WI38 human fibroblasts hOct4 levels using housekeeping hGAPDH (A) and hActin (B). (C and D) 3yr human fibroblasts hOct4 levels using housekeeping hGAPDH (C) and hActin (D). P-value ≤ 0.05; p-value ≤ 0.01**; p-value ≤ 0.001***.*

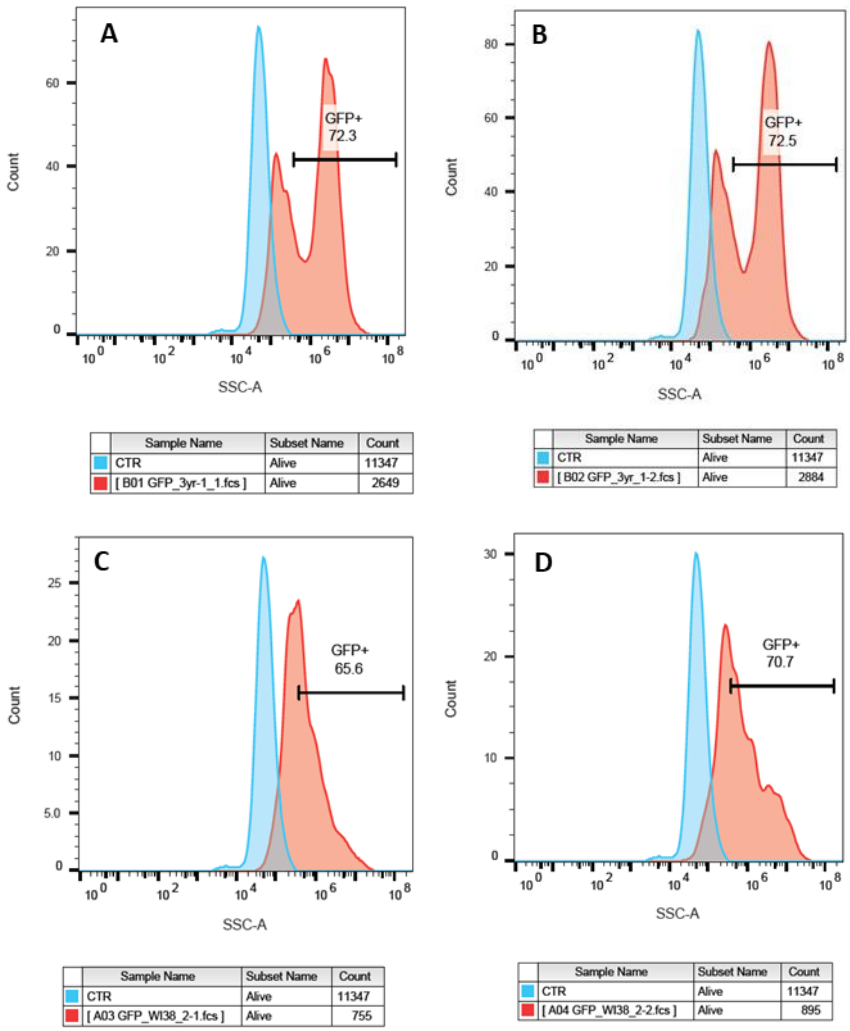


Figure 6.3 - Flow cytometry histogram results of 3yr and WI38 human fibroblasts expressing green fluoresnce plotting in red. 3yr and WI38 control plotted in blue.. (A) and (B) 72.3% and 72.5% of viable 3yr human fibroblasts expressing GFP, correspondently; mean= 72.4%; (C) and (D) 65.6% and 70.7% of viable WI38 human fibroblasts expressing GFP, correspondently, mean= 68.2%. Analysis performed using BD Accuri C6 with a cell count of 3000 for cells infected with GFP. Cell counting of control was set to 10000 cells.

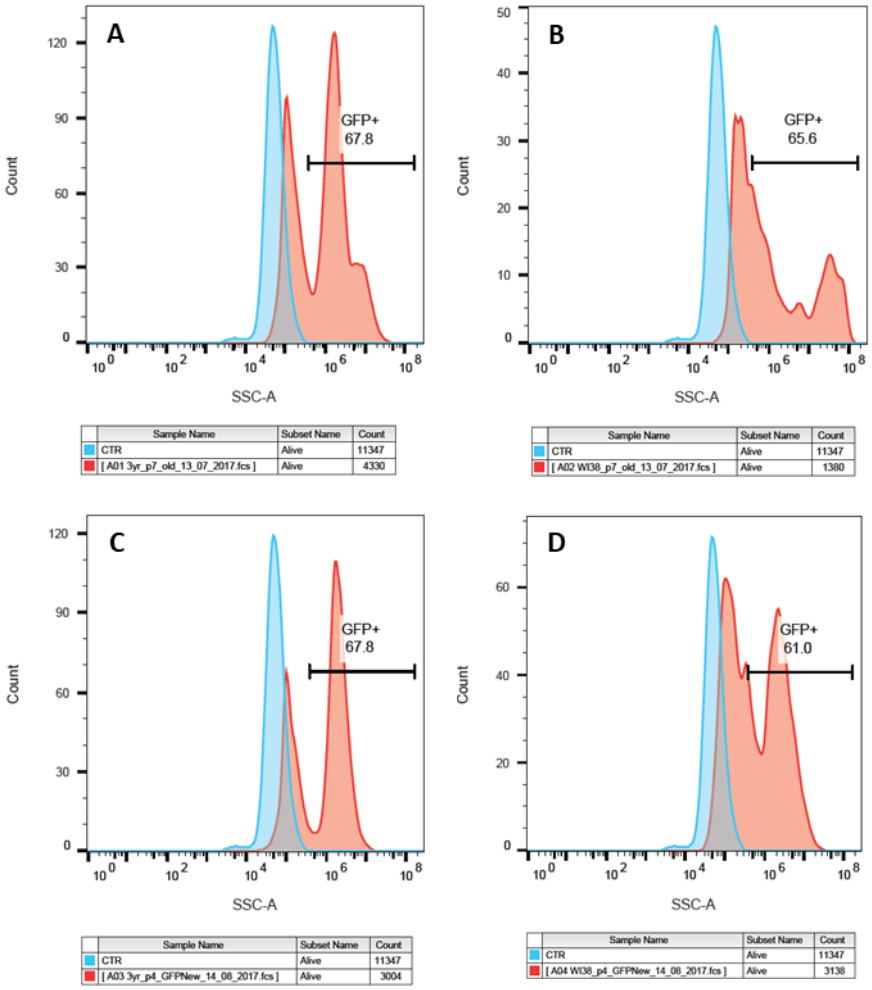


Figure 6.4 - Flow cytometry histogram results of 3yr and WI38 human fibroblasts with high (p7) passage expressing green fluoresnce plotting in red. 3yr and WI38 control plotted in blue. (A) 67.8% of viable 3yr human fibroblasts with a high passage (p7) expressing GFP. (B) 65.6% of viable WI38 human fibroblasts with a high passage (p7) expressing GFP. Analysis performed using BD Accuri C6 with a cell count of 3000 for cells infected with GFP. Cell counting of control was set to 10000 cells.

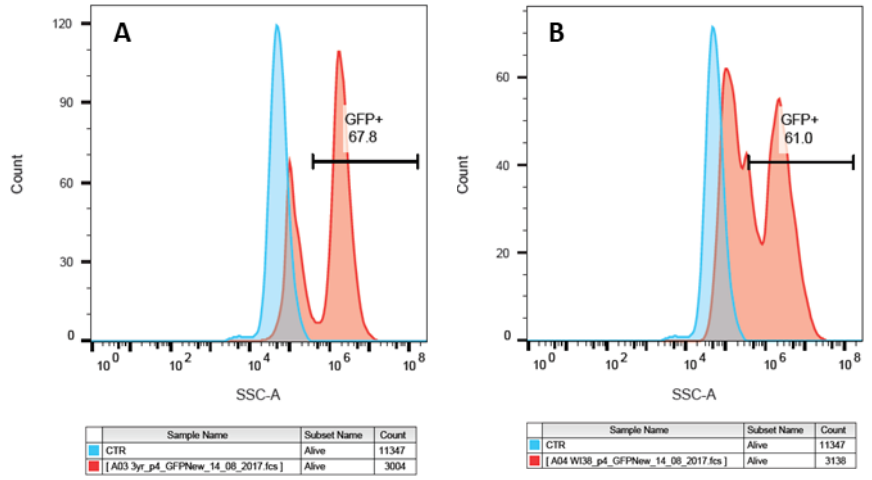


Figure 6.5 - Flow cytometry histogram results of 3yr and WI38 human fibroblasts with low (p4) passage expressing green fluoresnce plotting in red. 3yr and WI38 control plotted in blue. (A) 67.8% of vialbe 3yr human fibroblasts with a low passage (p4) expressing GFP. (B) 61.0% of vialbe WI38 human fibroblasts with a low passage (p4) expressing GFP. Analysis performed using BD Accuri C6 with a cell count of 3000 for cells infected with GFP. Cell counting of control was set to 10000 cells.

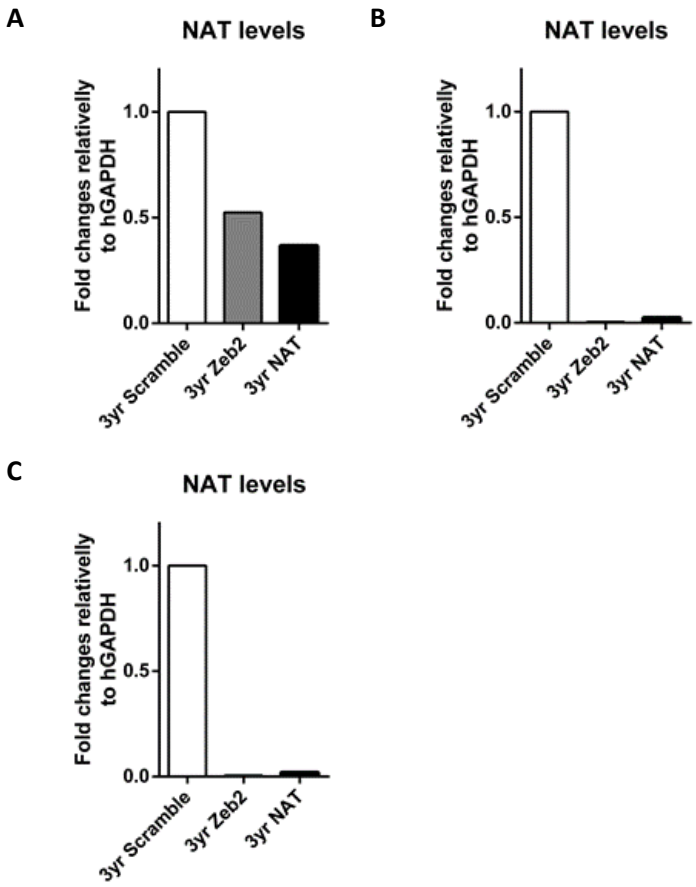


Figure 6.6 - Results of downregulation efficiency of hNAT expression level in 3yr human fibroblasts presented as $\Delta\Delta C_t$ normalized using non downregulation of lncRNAs Zeb2 and Zeb2NAT as control and HGAPDH and hActin as housekeeping gene. (A to C) ZEB2NAT expression levels in 3yr cells with knockdown of Zeb2 and hGAPDH as housekeeping.

Annex 2

NORAD gene sequence (29000097C17Rik)

5'-

ATAGTACCAGGACTATAAAGTGTTTTCCATAATAATAATGTCAGGGTGGTCCTGTACATTCCCTGCTTATAAGCC
TAGAATTAGCCTTTTTTCCAAGATCACCGACACTTTAGAAAAAAAGAATATTTTAGAAGCTTGGTCCTGACTTCT
GCTGTTGATGAAAGATGGCTTCGAATGGCCCATCATT**ACCTGGAAAGGCAAACACTG**TTGATCTTTCAGACCCCG
GGTGGATAACCTGGTACATCTAACCCAGGCCAGAGATTCAAAGTACCCTGACAAGCACCCCCACCCACCCACCCA
AGTGGAGGATAAAAATCTCTGCACTGCCACTTCAGATATTCCCTGAGCCTTACCTAGGTGAGTGGTGTTGCATCT
GACCTGTCCAAGCCTCTGGTTTGGACAATCACAAGGAAACTTCTTGCTTGGAAAACCGTGGCTCTTCAATGGGTG
CA***TGAACCAGCACGTAGCACAA***AGGAACAATCCAAGAAAAAAAGACACCACCTAAGCCCCTGCTCCCTGGCAGGC
AATGTGCTGTATCCTATAACTGCAGTGTAATTTGTGTGTGGCTTTGAAGACAACTGGCTCAGGTGGTAGAGAGGG
CTGCTGGGGAAAGACCACATTGTGGGGTGTCAGGTTCTTAC**ATTCAAGCCTGCAGAAGCAT**AAGGAACTGTGAGA
CCCCAGGTGAAATCTCACACTGCCTTGTCCAGGGAGGAAGCTAACCCCTAGGCAGGAAGGAAGGCACTGAAGATG
AGAGCCCCATCTATCTCAGCCCATTAGCATGAACACTTCGTTCTTTATTACACAAGCAGTAGAGACGCCCAAG

NORAD Gene >

TCTTTAAAGAAGCTTTTGTTTATTTTTTTAGTATATGCATTACAGGAAACAAGAAAGCAAGAACAAAGTCACCCA
GTCACAAGCAGCTCAGCTCGATGTATTGTTCTAGATCCTGCAAGCGTACACACACTTGGGTAATTGAGATTATAC
AGTATGTACCATGCTTTCCCCACATTGTGAATATTCACCAATTAAAAACCCCACACTACACAGGGATTTTGATAA
CCAAGAATACACCATGCCATCTCAATCTGTCAGACAAAATTGTAATCCTGTCTTCCTTGGCAACAGAAATTTCAT
AGAAGACAAAAATGGCAGCAGGCCTCTAGATAGATGTACTGGCAAAGTTACAATTAAATTCTGCATTCGCTCAGT
CTCAGCTTACAAGGAAACACAAAGCAACAGACCACACTAAGTATAGCATTCCTAAGCTTATCACCAGATCTATAT
ACTAGCATATGTAGAAGGGATCTTCCCACTCAACAAATGTATTCCACACGGAGCAGCTGGGAGACACACCTACGG
TAGGGGCTCTCAGGATGTCAGTCTCAACACACCTGCAGACCTCATCCTGTGAACCTGTGTCCCTCCAGTCAGCCT
AGCAACATGAATTTCCCCACACTGGCCAGACATGGCTTAGCAGTTCCATGTTCAAGAGGAACCCTTTTGTCAATT
CCACCATAGCCATTAGCTCTACTCTTAACCATGTGCTTTCCTTTAGGATGTAACATTCAGATGACAACTGTGCAT
AATGAACACAAGATTTTACCACGATGCACATAGGACCCCTGGGTAGAGCTGTCCCCTGGACAGAGGCACTAGCAG
CCTTTCCCTCAGCCATCCTCTTCACTGCACAACTCCTAAGTGCAGATGATGCTAGTGAGAGGATTAACAAAGGCC
ACACCAAGAAAACAGCCCTCCCCAAAGATGAGCCTGGCCTTTAGTGACTCTTGCTACTTCTAACATCCTTTAGGC
ACTGGGACTTACCTCTACTACCTGGCAACATTAAGTGAAATGCACACCCAACCAATGGCTAGACAACTCGAACCA
ACCAACAGCTACAGTGCCTCTTTCTGAGCACTAAGCTTTGCCCTAGCCCACCTCCATCACCTGGTAACAGCATCT
GCTCAAAGGCCAAGCTTAACCTCCATTCCCCCAGTGTGACCATTCCTGTAACAAAATCATCGTTAACAAAGTGGT
GTCAGCCCTCTATTTTTACATGTATGCAATACTGTATGCATGTGTGCACATGTGCCACTGCACTTGTGTGGAGGC
TAAAAGACAAGCTGTGAATCGCTTTTATCAAGTGGATCCCGGAAATCAAACTCAGGTCACCAGCTTAGTAGCAAG
CACCTTTACATGCTGAGCCACCCTACCAATCCGTTATGTACACTTTCTCACTGTTTTGAGACAAAGTCTATAGCC
CTGGTTGACCCAGAACAGAGATCCACCTGCCTCAGCCTAGAATTAAAGGCATGCATTACCACATCAGGTCTATCA
TGCACCTAAAAGGCTGGAGACCAGATGTTCCCAGTAAGCACTTTGGTTTCCCCAGCATGTGTACACCACAGTGAA
TACACTGCACACCTAACAAAGGCTGTGGTCAGACTGCCTTCTAAACAGCAAGGTTCTGGTCCAGCCCCTCTTCCA
GTTCTCTCCCCCCAGGGGCCATGTGTATGCATCTGTCATGTTTTTGTGGAATATTAACAATTTCACTTCCTAAGT
GGTTTTAAGGTTTTTTTCCTACAAATTTTAACATAATCTATAAATATGTCTTTTCACTAACTCTACTTTGTCACA
CACTGGCACCTTTACAGTCTAGGAAGACTGGGTGCTGCAAATTTGGACCTATGTTTTCTATACATTGCAAAGTAA
GATGCACAGATATAAGCAACAATGGATAACCTTGTATTACCAAACTTGCATCAGCATTAGGCTCTTCATTCATTT
TCTTCCTCCCTCCTAGACCTGCACAAGCAAGGCTGCATTGCTCAGAAGCAGTTTCTCTTTCTAGTTCCAAAAGAC
AACACTTCACCATGAGTTTAAACAGAATCTACAACACGTATTATTTACTATGTCTACTTTTAACATACTTTGTGC
ACTTCTAAACACCTAGAAGGACTAGATGTTTCAGATGAGGACATGAATTTGTCCCCTATATACATAGTAGTACTG
AAACCACACACATGTAACTAGTTAGATCTCCAAGTGTCTTCTGGTGAGGACATTCTGGTCTCCGACCTTTCTTCC
CCTTCAGCCCATCTTGTGTCCCGTGACATAGGCACGGGTGCCTATGTCACTCGTCACTTGGATGATACACTTTCA
CAACTCTTCAGAAGACAACCTTGCTATGAACTAACAGAGTATACAGAATGTGTGAGTTTACTAACTACTTTTGTC
ATACATTGGCAACCTCTTTAATACCTAGAGACTAGATATCATAGGACTCGTTTGTCCAGTATGTACAATATATAC
AGCATAGCAAAGTTACATGAGATGCACACAACACACAGGGGGATGGGTAAGCTAGAAATGTCTAGCACACAAATG
GAATCTTTTTGCAGTTCCTTCCCTCCCACCTCCTTGATCCACTGAACAAGCACAGCCGGTACACTGTTCTCAGAG
GCGGGCTAACAATTCTACTTCTACACACAGCATTTCTCCAGTTTAAAAAAACTATATACACAAAATATTATTCTA

TTACCTGTACCTTTAAATGTATCAAGCACTTTCAGATATCTAAAAGACTAGATGTTTCATGTAAGAACTTATCTG
CTATTGACATGGCACTGGTAAATAAAACTGTATGCAAAGCAATGGTTATTATCTGAGGTATCTTCTAAACATAAC
CATTTGGGCTTCGACCCACTTTCTTCTCCCTTCAAACCAATAGACAAATATAGGTATGTGTAATGCTTAGATACA
GCTGAGCAAATCCTATCCCCAAGCCATTTACAGAACAAAATTTTCTATGAATTTTAACGGTGTACACTATGTGCT
AAATTTACTACTTTGTCATACATTGGCAACCTCTTTAACATCTAGAGTCTAGATGTTGAAAATTTAGACATATTT
GTCCATTATATACAGTATACACACAGCAAAACAAAATGCAAGAATATAAAAATGGCATCTGAAAAGGTCCCATAT
ACACACACACTACATGGCCGTTGTCACTGCCTCCCAACACCTCCTCCAAACCCCGGAGCAAGTAAAGACAGTACT
GCTCAGAGGGCTCATCACCCCACTCCCCAAAGACAACGACTCATGAGTGTTTGCAAAAAGATATTTACAAAGTGA
TATTTCACTACCTCTACATTTAACATACGTTGGGCACTTCTAAACATCTAGATAGACCAGATGTTTCAAGTAAGG
ACTGTTGTCCACTATATACACAGCAGTCAGTAAATGGCACACGTAACAAGAGTAATTACAACTTTAACACAAAAT
GTAAAAAAAAAAAAAAAAAAAAGTGTTGTTCTACTAACTCTACTTTGTCATACACTGGCAACCTCTTTAACATCT
AGAGTGACTAGATGTTGTAAATTAGGACCCCTTTGTCCTTTACATACACTATGTACAAAGAAAAGCCAAACAAAA
TGGGATAGTGGGAGATAGTGGTTCCTTTCCACTCACGGACCACAGTAGCTTAGAGCTCTGGATTTCCTTCTGGGA
ACCAGGGCACAAACACAATGCGGTCAACTCAGGAGAGGTTTGGCTTTTCTTCTCCAAACAATAATTCAAATGAAT
CCTAACAAAAAGGTATTTACAGCATGTTTACTACTTATAAAATATGGCAGGCACTTCAGAGCAGCTAGAAAGACA
TTACCCCCAAAACCTGGCATTTTCAAACACACTCAAATACACACACACAGGCACACACGCAAACATGGTGTGGGG
GGTAAGAAATGAGGACTAAAATGCATACACAATACAATAGGATAAAAAAGAGAGACACGTCTTCTTCACACGACC
AATCTGGCTCCGGGCCTCTTCTTCCTCCTGCTGGATCCTCTAAGGTACTGAACAAACAGGGACGAGTGTCGCTCC
TGGGTCTTTCTAGAGGCGGCCTAAACAATTCCATTTCAAAATCACTTGCAGAAGACGTGTCCACGATCTTCTTTT
TATGGCGACTGGGTATTCACGTCATCTCCTACCTCTCCTTTACCCTGCGTCGGCAATCTCTACAGTTAGACGTCC
AGATTTAGCCAGAAGTCTAAAAAGGGTCGGGGGAAAGGTTGAGAGCAGCTTTTTCATATTATATACACAGGACTT
CTACGGCGACAGCGTTTCCGAATGGCCGCCTGCGGTTGTCGCGTCTCGGCTTCTCTCGGGTGGAGGGAAGACCCG
CGGCCCAAGAATCGATCGTTCCTCTCGCCGTGATTCGGGGCCCCGCTCACCTTCGTCCAGGCCCATCGAGGCCTC
GGTTCATCCAGGTCTGCGTGCTCGGCCACGGGTGGGCCTCCCGTCCCGAGCCGGCGATGGCAGAGTGTTCTTGCT
CAGACCCGCCGCCTTAGGCCGCAGAAGCCCCGAAGGCCCAGCCTTGGCCTTCAGGTGGATGCCATTCTTCCCGTC
ATGCCACGCTCGCCGCCACTTTCGAGGCGTCCTCGAAGGCCTGGCTGGGACTCCTCGGTGGGCCCGGCCCTGGAG
AGAGGGCGCCGGTGGAGTCGGGGAGGCCGCCGCCATGCAGGCGCATCGTTGGCAGGGCAAGAGTCGCTGCGGAAT
GGCCGCGTGGCCTAGCCGGTGGCTGCCTCGAGGGCGAGGGCTCCTGGCGTCCGCAGAGTAGGGACCGAGGCTGCT
CCCGGG***CCCCTCGGGCTCACCCGGC***GCGGGCCGAGGCCGTGAGGGGCCCGGGGCGAAGAAGGCCACACTCCCGAG
CTCGCTGCGCTCTCTGCCGGAACTGACTTGGCAATTCTTCTC

< NORAD gene

GCCCAGGCTGGAGCGGCACCGGAAGCCCGCAAAAACACCACCGGAGAGCTAACGTTTCCCCAACTGCAAGGTTCC
GGCCGGGAGCGGGGGGGGGGGGGGGGGGGGGAGCCTGGGGTTGGGGGCGGGGGTCTTTTTTTCTTTAAGTCTCT
TTACATCTTTTGTAGAGACT**AGAAGGCCAAGGGGAGAATA**AAAGGCCAAAAGTGAAAAAGAAAAAAAGAAAAAGC
CCACACTAGCAGTCCCAGCAATTTGTGATACATTATAAAAATAATTATAGGCAAACGCCTAAGGGCCTGAAGTAC
TACACTAGTCTCACTAGAAGTGGTGTTATAAACTAATTACCTAGTAAGATAGTGAACTGCTTGCCTCTTATCTAC
TCAAATACAACCCCATAGAATTTATTATTCCATGAGGCTAAGAAGGCAAATCGGACCCCTCCCCCCACATCACCC
CTCTCACAGTTTCTGTCCACGGTGGTGTGTGTGTGTGTGCGTGTGTCTGTCT– 3’

Cut

Primers